THE EFFECTS OF INTERPERSONAL GOAL, ACCOUNTABILITY AND TRAIT DESIRABILITY ON INFORMATION SEARCH STRATEGIES

By

MICHAEL F. WEIGOLD

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1989

ACKNOWLEDGMENTS

I wish to thank the members of my committee, Lawrence Severy, Mark Alicke, Richard Griggs, and John Lynch for tolerating my work habits and for their counsel. I am particularly indebted to my chairman, Dr. Barry Schlenker, from whom I have learned so much these past four years.

Thanks are due to my wife Mary Ann Ferguson for her support, editorial assistance, and willingness to sacrifice time and energy so that the dissertation could be completed on time. Thanks also to my parents who stressed the value of an education and whose sacrifices enabled each of their children to receive one.

TABLE OF CONTENTS

page

ACKNOWLEDGEMENTS.................................. ii

ABSTRACT.. iv

CHAPTERS

1 INTRODUCTION.................................... 1

Social Information-Gathering Strategies.......... 1
Factors Likely to Affect Information Search...... 41
Overview... 49

2 METHOD.. 52

Subjects... 52
Procedure.. 52
Coding... 56

3 RESULTS... 58

Design and Analyses.............................. 58
Primary Analyses................................. 63
Additional Analyses.............................. 73

4 DISCUSSION...................................... 80

What Was Thought Would Happen and Did............ 81
What Happened but Wasn't Expected................ 84
What Was Thought Would Happen but Didn't......... 88
Information Search Strategies: Future Directions... 91

APPENDICES

A MANIPULATIONS................................... 97

B DEPENDENT VARIABLES............................. 103

REFERENCES.. 108

BIOGRAPHICAL SKETCH............................... 119

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

THE EFFECTS OF INTERPERSONAL GOAL, ACCOUNTABILITY AND TRAIT DESIRABILITY ON INFORMATION SEARCH STRATEGIES

By

Michael F. Weigold

August, 1989

Chairman: Dr. Barry Schlenker
Major Department: Psychology

Research investigating social information-gathering strategies has yielded two formulations, the confirmatory and diagnosticity models. A review of the models suggests important problems with each, and a new framework is proposed. An experiment is described which tests the impact of accountability, trait desirability, and interpersonal goal on interpersonal search strategies. Two hundred and thirty-seven subjects composed question lists to test whether or not a target possessed one of four traits. Lists were evaluated by judges to determine the quantity, quality, and consistency of the questions. Findings pose problems for both the consistency and diagnosticity models. The consistency formulation is challenged by the relative scarcity of questions that can be logically evaluated for hypothesis consistency, and strong preference of subjects assigned to test for a negative trait for questions inconsistent with the hypothesis. The diagnosticity formulation is challenged by the systematic differences in question diagnosticity across interview conditions and within question types. Implications of the research are discussed.

CHAPTER 1
INTRODUCTION

Social Information-Gathering Strategies

How do people obtain information to test a hypothesis about another individual? While there is a volumous literature in social psychology dealing with social judgment, much of it only touches on issues relevant to information acquisition. For example, many studies of impression formation examine how social cognitions vary as a function of subjects' exposure to experimentally controlled stimuli or contexts (Asch, 1946; Doherty, Weigold, & Schlenker, 1989; Doherty, Weigold, & Schlenker, in press; Higgins & Stangor, 1988; Jones & Harris, 1967; Ross, 1977; Ross, Amabile, & Steinmentz, 1977; Schlenker, Hallam, & McCown, 1983; Schlenker & Leary, 1982; Swann, 1984). The benefits associated with precise control over the social stimulus come at the potential cost that subjects be relatively passive information processors. Subjects must use the information given, whether or not such information would be used or sought after in more typical decision environments (Swann, 1984).

Recently, however, psychologists have acknowledged that a passive view of persons may be inadequate. As the influential metaphors of "man as scientist" proposed by George Kelly (1955) or reciprocal determinism as proposed by Bandura (1987) suggest, people are active participants in their social environments. Thus, many of the important dynamics associated with impression formation and social learning may not operate the same way when people are presented with a set of information as they do when people must seek out such information for themselves.

Recent trends in social research reflect the idea that goal-directed activities instigate processes that are poorly revealed in standard experimental paradigms. For example, in the attempt to understand biases in impression formation and stereotyping, social psychologists have begun investigating how people acquire information for making a social judgment. This work derives, in part, from both studies of the information people prefer for testing hypotheses (Griggs & Cox, 1982, 1983; Wason, 1960, 1968; Wason & Johnson-Laird, 1972) and social psychological research in selective exposure to information (Festinger, 1957; Frey, 1986).

Perhaps because of its link to Wason's research, some social psychologists have labeled strategies of social information acquisition "hypothesis testing." The term is somewhat misleading, however, since research of this kind typically does not investigate decision rules for accepting or rejecting hypotheses. Rather, studies as typified by

Snyder & Swann (1978) concern themselves with how social perceivers _gather_ or _access information_ for purposes of testing a hypothesis (Skov & Sherman, 1986; Swann & Guliano, 1987; Trope & Bassok, 1984), whether from the environment (Trope, Bassok, & Alon, 1984) or from memory (Snyder & Cantor, 1979). Hence, to avoid ambiguity, the procedures used by people to gather information about a target for purposes of testing a hypothesis or making a social judgment are labeled _social information-gathering strategies_.

Knowledge about social information gathering strategies may be important for several reasons. First, the means by which people acquire social information should affect the accuracy of resulting judgments and inferences. Systematic departures from "optimal" methods may distort the conclusions people reach about targets (Skov & Sherman, 1986; Snyder, 1984; Snyder & Swann, 1978; Swann, 1984). For example, biased search for information that supports an erroneous hypothesis may result in the perseverance of unfounded beliefs or superstitions (Fischoff & Beyeth-Marom, 1983; Janis & Mann, 1977; Snyder & Swann, 1978; Skov & Sherman, 1986). Evidence that people prefer information gathering strategies that consistently bias impression formation may thus have relevance for issues of prejudice and stereotyping.

Second, examining information-seeking strategies can shed light on what kinds of information people most prefer for making judgments. At least three "preferred" kinds of information have been reported in research examining self-relevant feedback: 1) information that supports the subject's working hypotheses or deeply-held beliefs (Swann, 1983, 1985; Swann, Griffin, Predmore, & Gaines, 1987), 2) information that maintains the subject's self-esteem (Mettee & Aronson, 1974; Schlenker & Weigold, 1989; Shrauger, 1975), and 3) information of maximum diagnosticity (Alicke, 1987; Trope, 1983, 1986). As will be discussed shortly, there is at present a debate concerning the kinds of information people prefer for testing a hypothesis about another person.

Third, people's preferences for certain kinds of information may provide valuable information about structures and processes relevant to more general cognitive issues, such as social memory (Hastie & Park, 1986) and social judgment (Swann, 1984). For example, preferences for acquiring certain kinds of information may in part be due to the ease with which such information can be compared to information stored in memory (Fischoff & Beyth-Marom, 1983; Snyder & Campbell, 1980), or to the ease with which such information can be analyzed for purposes of determining covariation (Alison & Messick, 1988; Trope, Bassok, & Alon, 1984).

Finally, the strategies of information seeking that people use can provide insight into the unique dynamics of interpersonal relations. Social information seeking can be differentiated from non-social information seeking in that it often takes place a context of interpersonal interaction. Thus, rules which may hold true in situations of minimal social pressures may not generalize to more complex social environments. Both Tetlock (1985) and Schlenker (1986b; Schlenker & Weigold, 1989) review literatures suggesting that decision processes and outputs observed in nonsocial contexts are often affected or changed in contexts characterized by social pressures and accountability. Therefore, a social psychological analysis of information seeking strategies may help to delineate the boundaries of effects obtained in nonsocial environment.

Since Snyder and Swann's (1978) provocative series of experiments, a debate has centered on whether people favor one of two rather broadly defined strategies for acquiring information. Snyder and Swann conclude that their experiments demonstrate a pervasive preference on the part of individuals for information that confirms their working hypotheses. They further suggest that this preference can lead to an information-gathering process that elicits hypothesis-confirming evidence from target others, thereby creating a self-fulfilling prophecy (Rosenthal, 1971; Snyder, 1984). If true, such processes may be implicated in the creation and perpetuation of ill-founded stereotypes and prejudices. Snyder and Swann's hypothesis that people preferentially solicit hypothesis-confirming evidence from targets is labeled the confirmatory-strategy hypothesis.

In contrast, Trope and Bassok (1982) argue that individuals prefer diagnostic information for testing a hypothesis or making a social judgment. Diagnostic information is that information which best permits the hypothesis tester to discriminate between the hypothesis under consideration and alternative hypotheses (Bassok & Trope, 1984; Kruglanski & Mayseless, 1988; Skov & Sherman, 1986; Trope & Bassok, 1982; Trope, Bassok, & Alon, 1984).

Trope and Bassok argue that Snyder and Swann's research findings are based on flawed methodology that unfairly restricts subjects' ability to seek the information they would most prefer. They also question whether ordinary information-seeking strategies bias subsequent judgments of the target. The claim of Trope and Bassok that laypersons most prefer information-acquisition strategies that are maximumally diagnostic with respect to the hypothesis and its alternatives is called the diagnostic-strategy hypothesis.

In the following sections, greater detail is provided about each of these hypotheses and their supporting bodies of evidence. Next, an alternative viewpoint is proposed and constructs that are believed to mediate information-search strategies are presented. In the concluding section of the chapter, hypotheses for the present research are offered.

The Confirmatory-Strategy Hypothesis

The first strategy endorsed as "pervasive" is the confirmatory strategy (Snyder, 1981; Snyder & Gangstead, 1981; Snyder & Swann, 1978; Swann & Guliano, 1987). In their initial investigation, Snyder and Swann (1978) suggested people might use one of three strategies in acquiring social information:

> In one strategy, the individual might preferentially search for behavioral evidence that would tend to confirm the hypothesis under scrutiny . . . . In another strategy, the individual might preferentially search for evidence that would disconfirm the hypothesis . . . . In a third strategy, the individual might search for hypothesis-confirming and hypothesis-disconfirming evidence with equal diligence. (p. 1203)

In study one, subjects were to discover whether a person in another room had a personality matching a description outlined on card. The descriptions were those of a "prototypical" introvert or extravert. For example, the extravert description read "Extraverts are typically outgoing, sociable, energetic, confident, talkative, and enthusiastic." The introvert description read in part that "Introverts are typically shy, timid, reserved, quiet, distant, and retiring." In addition, Snyder and Swann manipulated the degree subjects could believe that the hypothesis described the target. High certainty subjects were told that the profile reflected the results of a personality test completed by the target. Low certainty subjects were not provided any reason to believe that the hypothesis necessarily described the target. Subjects then selected from a list of 26 questions the 12 they most preferred for making their judgment. Eleven of the 26 questions had been classified as ones "that the majority of the rater-judges thought would typically be asked of people _already_ _known_ to be extraverts" (p. 1205). Ten questions were judged to be typical of those asked of introverts. Five questions were judged to be neutral in that either "there was no consensus that they were extraverted questions or introverted questions" or they were "classified by the majority of the rater-judges as irrelevant" to the traits (p. 1205).

Results showed that individuals assigned the extrovert hypothesis preferred to ask more extravert-confirming questions than did persons assigned the introvert hypothesis. Additionally, those given the introvert hypothesis preferred more introvert-confirming questions than those induced to test the extravert hypothesis. Furthermore, the choice of questions was not qualified by the certainty manipulation; both high and low certainty subjects showed preferences for "confirmatory" questions.

In a second experiment, Snyder and Swann assigned 80 female undergraduates to roles as either interviewers or targets. Interviewers were given instructions identical to those of subjects in the low-certainty conditions of the first study, with the exception that they actually asked their preferred questions to a target. The interviewer-target interactions were recorded and edited so that 12 judges could be exposed to the first and last three answers provided by targets. After listening to the tape, judges evaluated the target on 10 dimensions relevant to introversion-extraversion: talkative-quiet, unsociable-sociable, friendly-unfriendly, poised-awkward, introverted-extraverted, enthusiastic-apathetic, shy-outgoing, energetic-relaxed, cold-warm, unconfident-confident.

Results again indicated that subjects testing a particular hypothesis were more likely to select questions consistent with that hypothesis than were subjects testing the alternative. More importantly, ratings of the target appeared to be affected by the assigned hypothesis. Significant differences emerged for four of the ten dimensions and two more dimensions showed marginally significant differences. In addition, the overall MANOVA for the 10 items achieved significance. Means for the target ratings showed that judges were more likely to perceive targets' personalities reflected the hypothesized versus alternative trait.

In two additional experiments, Snyder and Swann attempted to discover boundary conditions on the confirmatory phenomenon. In experiment three they varied the base-rate of for persons possessing the trait in a group from which the target was supposedly randomly sampled. In experiment four they varied motivation for accuracy by offering some subjects a cash prize for accurately testing the hypothesis. Neither of the manipulations attenuated subjects' preference for confirmatory questions in testing a hypothesis about another person.

Snyder and Swann concluded that "the structure and process of human thought fosters and promotes the ready and willing adoption of confirmatory strategies . . . ." (p. 1210). Evidence confirming the existence of these structures and processes has been shown in people's preference for positive versus negative instances of a concept (Hovland & Weis, 1953), the more powerful impact of confirming versus disconfirming instances on inductive conclusions (Golub, Rossman, & Abelson, 1973), positive versus negative associations in judgments of covariation (Jenkins & Ward, 1965), and similar versus distinctive features in judgments of similarity (Tversky, 1977). Snyder and Swann argue that when people anticipate having to acquire information for testing a hypothesis they construct cognitive scenarios of the upcoming interaction. Imagining this construction, according to the researchers, is easier when targets are pictured acting in accord with rather than in opposition to the hypothesis.

Furthermore, processes of confirmation in turn have the unfortunate result of leading people to sample data

> in which hypothesis confirming evidence will be overrepresented and hypothesis-disconfirming evidence will be underrepresented. For there is every reason to believe that most people, as targets, will be 'generous' in providing specific instances of hypothesis-confirming actions. (p. 1211)

Snyder and Swann appear to reject the idea that people will use disconfirmatory or equal-opportunity strategies for testing hypotheses. While recognizing that further evidence was necessary for justifying broad conclusions, they note that

> At the very least, we can assert that we have yet to identify any procedure that will induce individuals to eschew [confirmatory] strategies in favor of either disconfirmatory or equal opportunity strategies. (p. 1210).

While offering cautions about the generalizability of their findings, Snyder and Swann describe the potential importance of strict reliance on strategies which seek only confirming information:

> From this perspective, it becomes easier to understand why so many popular beliefs about other people (in particular, clearly erroneous social and cultural stereotypes) are so stubbornly resistant to change. Even if one were to develop sufficient doubt about the accuracy of these beliefs to proceed to test them actively, one nevertheless might be likely to 'find' all the evidence one needs to confirm and retain these beliefs. (p. 1212)

Snyder and Cantor (1979) investigated whether the confirmatory bias also operates when people must recall evidence to test a hypothesis. Subjects in their first study read information about the activities and interests of a purportedly real person named Jane. The information was designed so that half of Jane's activities were characteristic of introverts and half were characteristic of extraverts.

Two days later, half the subjects assessed Jane's suitability for the "rather extraverted position of real estate salesperson" while the other half evaluated her suitability for the "rather introverted position of research librarian." Thumbnail sketches provided prototypical attributes of the relevant trait. Also varied was whether or not Jane was merely considering the position in question, or whether she had already accepted the job. Subjects spent 7 minutes attempting to recall facts from the story that were relevant to the suitability decision and then rated Jane's suitability for the job.

Judges rated the recalled facts in terms of whether they were introvert-relevant or extravert-relevant. The ratings were then rescored to reflect whether the recalled facts for each subject were hypothesis-confirming or hypothesis-disconfirming. Results showed the predicted main effect for subjects remembering more hypothesis-confirming rather than hypothesis-disconfirming facts. However, an interaction of the hypothesis and whether or not the job was described as taken versus being considered showed that subjects only remembered more hypothesis-confirming facts when told that Jane was considering the job. There were no differences in confirming versus disconfirming facts recalled for subjects told Jane had already taken the job. Additionally, ratings of Jane's suitability for the job were not significantly affected by either of the manipulations. Snyder and Cantor explained the failure to find differences in the "taken" condition by saying

> Participants in the Taken conditions knew that both Jane and her employer had already judged her to be at least minimally suited for her job. Thus, the Taken conditions may have contained presumptive evidence of job suitability. Consequently, participants in the Taken conditions may have believed that less confirming evidence was necessary to accept the hypothesis--and thus accumulated less confirming evidence before accepting the hypothesis--than did participants in the Apply conditions. (p. 336)

They fail to note, however, that such an explanation contradicts the findings of Snyder and Swann (1978) that confidence in the hypothesis did not affect information search.

In a second experiment, Snyder and Cantor replicated the "apply" conditions of study one, however this time thumbnail sketches of the trait prototypes were not included and subjects rated Jane's suitability both for the hypothesis-congruent job and for its alternative. The results showed that subjects again recalled more hypothesis-confirming facts, and additionally rated Jane as better suited for the hypothesis-confirming than the hypothesis-disconfirming position. They concluded that confirmation biases are present not only when people acquire information from the environment but also during information searches in memory.

Snyder and Campbell (1980) sought to determine whether confirmatory information seeking resulted from the way subjects frame the initial hypothesis. The typical stimuli used in information search experiments are explicit about the attributes that _confirm_ membership in the category, but are mute on attributes that _disconfirm_ membership in the category. Thus, Snyder and Campbell framed hypotheses for some subjects in terms of attributes that would confirm _and_ attributes that would disconfirm the hypothesis. Other subjects received hypotheses only in terms of attributes that would confirm the hypothesis. The procedure otherwise mirrored that of Snyder and Swann (1978). Results showed that subjects once again showed a preference for questions that would confirm the hypothesis. In addition, the framing manipulation did not affect these preferences.

Swann and Guliano (1987) sought to determine whether people spontaneously engage in confirmatory search strategies for hypothesis-relevant information. In their first experiment, subjects were instructed to prepare questions for determining whether a hypothetical respondent "possessed the characteristics described in a trait profile." Each subject generated five questions for each of two profiles. The first profile described either dominance or submissiveness, while the second described either extraversion or introversion. Judges later rated the "characteristics of a person who would respond 'yes' to a question (e.g., extraverted, submissive, etc.)." The rated questions were assigned to one of four categories: (1) confirmatory questions, where a "yes" answer indicated the presence of the trait, (2) disconfirmatory questions, where a "yes" would indicate presence of the opposite trait, (3) equal-opportunity questions, where respondents could provide answers that could either exemplify the trait or its opposite, and (4) irrelevant questions, or those that bore no resemblance to the trait or its opposite.

Results showed that 30 percent of the responses fell into one or the other of the latter two categories (exact proportions of each are not mentioned). These were subsequently dropped from further analysis, so that the reported results dealt with only the questions that could be classified as confirmatory or disconfirmatory. Results showed that subjects were again more likely to ask questions for which a "yes" response was consistent with their hypothesized trait than were subjects testing the alternative trait.

Study 2 attempted to determine conditions under which subjects would prefer "highly constraining" to "mildly constraining" questions. According to Swann & Guliano, highly constraining questions are those that severely limit the degree to which respondents can provide evidence inconsistent with the hypothesis. For example, a constraining question for extraversion would be (from Snyder and Swann, 1978): "What would you do if you wanted to liven things up at a party?" Answers to this question, irrespective of the respondent's true level of extraversion, would be unlikely to falsify a hypothesis that the respondent is an extravert. Mildly constraining questions are those for which a "yes" answer provides evidence consistent with the hypothesis, but which does not restrict the target from providing evidence falsifying the hypothesis, for example (from Swann & Guliano): "Do you like to entertain people?"

Swann and Guliano propose that when people hold their hypotheses with high versus low confidence, they are more likely to prefer constraining questions. Subjects were given either good or poor reason to believe a target was an introvert or extravert. They then choose eight questions from a list of 20 that for the target to answer. The list was constructed of the following types of questions: five confirmatory and mildly constraining, five confirmatory and highly constraining, five disconfirmatory and mildly constraining, and five disconfirmatory and highly constraining.

Results showed a marginally significant preference for confirmatory questions ($\underline{p} < .06$), qualified by a question type x belief certainty interaction demonstrating that high but not low certainty subjects preferred confirmatory questions. In addition, low certainty subjects did not differ in preference for mildly versus highly constraining questions, while high certainty subjects preferred the latter ($p < .05$).

In a final study, Swann and Guliano asked undergraduates to rate "how much they could learn about a respondent from listening to him or her answer" each of six questions, all taken from Snyder & Swann (1978). Two of the questions were relevant to introversion, two to extraversion, and two neutral with respect to the traits. The results of planned comparisons pitting the confirmatory questions against the disconfirmatory and neutral suggested confirmatory questions were viewed as more informative than other types.

Summary: The Confirmatory Formulation

The research program of Mark Snyder, William Swann, and their colleagues suggests individuals acquire information for purposes of testing a hypothesis about a target by formulating questions that are consistent with the hypothesis. Despite a series of empirical studies that are presented as supporting the consistency position, some problems remain. For example, "consistent" questions appear at various times to mean: (a) questions for which an affirmative response indicates the presence of the hypothesized trait and (b) questions biased towards eliciting evidence consistent with the target's possession of the hypothesized trait. As will be elaborated later, these two meanings are quite different. There are several other difficulties with Snyder and Swann's position, which are discussed under the diagnosticity formulation.

Early research (Snyder & Campbell, 1980; Snyder & Cantor, 1979; Snyder & Swann, 1978) suggested that people's use of a consistency strategy is subject to few mediating conditions. However, later research (Swann & Guliano, 1987) suggests confirmatory strategies are most likely when questioners have high as opposed to low confidence in their hypotheses. Snyder and his colleagues argue that gathering information in a hypothesis consistent fashion is pervasive, relatively nonconscious (i.e., people do not attempt deliberately to confirm their hypotheses), and has potentially deleterious effects (reinforcing negative or erroneous expectations about the target).

The Diagnostic-Strategy Hypothesis

A second strategy proposed as a model for the information gathering activities of social judges is the diagnostic strategy. Proponents of this approach argue that individuals most prefer information that can best discriminate between a working hypothesis and its alternative(s). An early study reporting evidence of subjects' preference for diagnostic information is reported by Trope and Bassok (1982). Subjects were told that the study concerned "the issue of handwriting for determining personality traits." According to subjects' instructions, people could be reliably classified as either those who are predominantly analytical or those who are predominantly intuitive. Half of the subjects were to determine whether a target was analytical, half were to discover if the target was intuitive.

Subjects indicated their preferences for handwriting features that could help discriminate between two traits. The experiment orthogonally varied the conditional probabilities of each feature given a target's possession of a trait (very low, moderately low, moderately high, very high) and the diagnosticity of each feature (low, high). Trope and Bassok argue that the confirmatory bias hypothesis proposed by Snyder and Swann predicts subjects will attend to the conditional probabilities but ignore diagnosticity. When the feature is likely given the trait, subjects should prefer the feature irrespective of how well it discriminates the hypothesis under question and its alternative. In contrast, Trope and Bassok proposed that subjects would be most concerned with the diagnosticity of the feature irrespective of its conditional probability.

The diagnosticity of a feature is defined in terms of its capacity to distinguish the hypothesis from its alternative. A measure of diagnosticity, $\underline{D}$, is calculated by multiplying the probability of the feature, P(F), by the likelihood ratio of the feature, LR(F), then dividing by the probability of the feature's absence, P(F), multiplied by the likelihood ratio of the feature's absence, LR(F).

$$D = P(F) \times LR(F) / P(F) \times LR(F)$$

The likelihood ratio of the feature's presence is the conditional probability of the feature given the hypothesized trait, P(F/H), divided by the conditional probability of the feature given the alternative trait, P(F/A). Conversely, the likelihood ratio of the feature's absence is the conditional probability that the feature will be absent given the trait, P(F/H), divided by the conditional probability of the feature's absence given the alternative, P(F/A).

In the Trope and Bassok experiment, low-diagnostic features had *D*s of 1.25, while features high in diagnosticity had *D*s of 3.75. The conditional probabilities of the features for each of the hypotheses were presented to subjects in the form of bar graphs. Subjects rated their preferences for each feature on a scale ranging from "not at all interested" to "very interested." A check of the diagnosticity rating was obtained by asking a separate set of subjects to rate the features in terms of how well they discriminated the groups. The results of the study were highly supportive of the diagnosticity hypothesis. Subjects significantly preferred features that were high versus low in diagnosticity, but showed no preference for features with higher conditional probabilities given the hypothesis.

In a second study, Trope and Bassok controlled for the overall probability of each feature across both traits. In study one this had tended to be higher for diagnostic features. Additionally, they manipulated three rather than two levels of diagnosticity. The results again supported the diagnosticity hypothesis, with more diagnostic features being preferred over less diagnostic ones. No coherent pattern emerged for the probability of the feature given the hypothesis.

A third study extended the generalizability of previous findings by manipulating the diagnosticity of continuous features, that is, features with a range of values for each of the hypotheses. For continuous features, diagnosticity is a negative function of the degree to which the features' conditional-probability distributions overlap. The nonoverlapping interval for each conditional distribution may also be varied independently of the nonoverlapping interval for the alternative, thus permitting varying degrees of conditional diagnosticities.

Trope and Bassok argued that in Snyder and Swann's framework features should be most preferred when they are diagnostic with regard to the hypothesized trait but undiagnostic with regard to the alternative. In contrast, the diagnosticity hypothesis predicts people will most prefer features that are maximally diagnostic across both traits. Total diagnosticity is an additive function of a feature's conditional diagnosticities for each trait.

The results again supported the diagnosticity hypothesis. Subjects most preferred features that were diagnostic for both the hypothesized trait and the alternative. In addition, subjects did not prefer features that were more diagnostic of the hypothesis and less of the alternative to those that were more diagnostic of the alternative and less of the hypothesis, as Snyder and Swann would predict. Moreover, questions that could only confirm the hypothesis were not preferred to ones that could only disconfirm it.

Bassok and Trope (1984) tested whether the relative salience of the hypothesis and its alternative would influence the strategies people used to gather information. In one condition subjects were presented with a detailed profile of the hypothesized trait. In a second condition, subjects were presented with detailed profiles of both the hypothesized trait and its alternative. Finally, subjects in a third condition received only a profile of the alternative trait.

From a list of 24 "yes-no" questions, subjects were instructed to select 12 that would allow them to test whether a target was an extravert or introvert. Questions were generated by a pretest group of subjects and selected on the basis of judges' ratings. The final list of questions contained items from all four combinations of high or low diagnosticity with extraversion or introversion. A salience manipulation consisted of giving subjects the appropriate personality description of introverts and/or extraverts used by Snyder and Swann (1978).

For extraverts, Bassok and Trope's instructions read:

> You will be presented with a list of 24 questions. Your task is to select 12 questions that will help you determine whether any given student tends to extraversion. Twenty Hebrew University students will answer the questions you will select. Ten of the students were classified by a panel of psychologists as tending to extraversion. The efficiency of your questions will be determined by their ability to determine whether each of the students tends to extraversion. (p. 206)

The appropriate words were substituted for the introvert hypothesis conditions. Results showed that across all conditions, subjects preferred questions of high versus low diagnosticity ($p < .0001$), and that question diagnosticity accounted for 86 percent of systematic variance. Within a particular hypothesis, subjects tended to prefer questions consistent with the hypothesis versus its alternative ($p < .001$). However, subjects consistently preferred questions about the alternative that were more diagnostic than questions about the hypothesis. In comparing the proportions of questions selected, the authors note that a slight majority of questions were consistent with the hypothesis (56 percent) while a substantial majority were of high diagnosticity (71 percent). In addition, a preference for questions consistent with the hypothesis was significant only when the profile of the hypothesis was given alone, and not for the other two conditions.

In a follow-up, Bassok and Trope had subjects choose six questions from a list of 12 to test a hypothesis about introversion or extraversion. All 12 questions in the list were of high diagnosticity, with six confirmatory for introversion and six confirmatory for extraversion. Results revealed a significant assigned hypothesis by preferred question type interaction showing that subjects were slightly more likely to favor hypothesis-consistent questions. However, the degree of overall preference for consistent questions was small (56 percent), and the authors conclusions were that "people are primarily guided by diagnosticity considerations in gathering information for testing their hypotheses" (p. 212).

Trope, Bassok, and Alon (1984) argued that the questions used by Snyder and Swann (1978) were biased, in that they "assumed the interviewee was either an introvert or an extrovert, thus making it difficult for him or her to express the opposite trait" (p. 92). The study tested whether or not people who spontaneously generate questions for purposes of testing a hypothesis would generate similarly constraining (and hence undiagnostic) questions, or whether people would instead generate diagnostic questions.

The authors suggest that subjects can ask four kinds of diagnostic questions: (1) consistent questions, or "questions about features that believed to be associated with the hypothesized trait," (2) inconsistent questions, or those concerning features "believed to be associated with the alternative trait," (3) bidirectional questions, or those that "present the respondent with a choice between features that are associated with the hypothesis and those that are associated with the alternative," and (4) open-ended questions, or questions that "refer to a set of features, some of which may be consistent with the hypothesized trait and some with the alternative trait" (p. 93). These exhibit close correspondence to Swann and Guliano's (1987) classifications of (respectively): confirmatory, disconfirmatory, equal-opportunity, and open-ended questions.

Trope et al. (1984) claimed that the confirmatory-bias hypothesis proposes that people prefer to ask consistent questions, or questions about features associated with the to-be-tested trait. However, the authors also note that such a strategy is not "biased" if the questions are diagnostic, because targets can respond with evidence supporting or challenging the hypothesis. Hence, the "bias" in the research of Snyder and Swann lies not in the strategies preferred by subjects but in the questions subjects are constrained to choose from. Trope et al. acknowledge that people may prefer consistent (as they define the term) questions, because:

> . . . it may be easier to think about confirmation in terms of affirmation, i.e., the _presence_ of consistent features, then in terms of negation, i.e., the _absence_ of inconsistent features; and it may be easier to think about disconfirmation in terms of negation, i.e., the _absence_ of consistent features, then in terms of affirmation, i.e., the _presence_ of inconsistent features. (p. 93)

Subjects in the experiment were told to formulate questions for use in an upcoming interview with a respondent. Half the subjects were to attempt to determine if the target was introverted, while half were to determine if the target was extraverted. Each subject received the appropriate personality description of introverts or extraverts used in Snyder and Swann (1978) and generated a minimum of four questions.

Two judges examined the trait descriptions and evaluated randomly-ordered lists of questions generated by the subjects. In addition, judges were given the identical questions used by Snyder and Swann (1978) without knowledge that these had not been generated by subjects. Judges were to classify questions into one of five categories: 1) biased extrovert, 2) biased introvert, 3) extrovert, 4) introvert, 5) bidirectional and open-ended. Biased questions were considered to be those that asked about characteristics of the a trait without permitting evidence indicative of the alternative.

Results showed that a) none of the subject-generated questions were judged to be biased, and b) all of the extrovert/ introvert questions taken from Snyder and Swann (1978) were judged as biased. The only questions used by Snyder and Swann that were not rated as biased were the neutral questions. For testing a hypothesis, subjects overwhelmingly preferred to ask open-ended and bidirectional questions (73 percent). Furthermore, the number of questions subjects generated that were consistent with their hypotheses were not significantly greater than the number generated that were inconsistent with their hypotheses (15 percent to 12 percent, respectively). In fact, whether testing for introversion or extroversion, subjects preferred to ask questions consistent with extroversion (20 percent) to those consistent with introversion (7 percent).

In a second experiment, Trope et al. (1984) again told subjects to formulate questions for an anticipated interview with a person for purposes of determining if the target was either an introvert or an extrovert. A third group of subjects was not assigned a hypothesis, but was told to generate questions that could discriminate introverts and extroverts. The groups were again given the Snyder and Swann (1978) trait descriptions (the control was given both descriptions). Questions were again coded by judges into the categories used in study 1. The diagnosticity of each question was also rated by 15 independent judges who read the trait descriptions and were told:

> Your task is to assess the extent to which each question can discriminate between an extroverted student and an introverted student. Try to imagine the answers interviewees may give to each question. To the extent that an extroverted student and an introverted student are likely to give different answers you should rate the questions as discriminating well between the two types of people. To the extent that an extroverted student and an introverted student are likely to give similar answers you should rate the questions as not discriminating between these two types of people. (p. 95)

Of the 273 questions generated by subjects, only two were judged to be biased. Once again, the majority were either bidirectional (43 percent) or open-ended (22 percent). These types, which are difficult to view as biased since do not favor responses consistent with or inconsistent with the hypothesized trait, were no less preferred by those testing versus not testing a hypothesis. The groups also did not differ in the judged diagnosticity of their questions.

Summary: The Diagnosticity Formulation

The diagnosticity hypothesis of Trope, Bassok, and colleagues proposes that people most prefer information that discriminates between a hypothesis and its alternatives. In a series of experiments pitting diagnosticity manipulations against consistency manipulations, the researchers have shown that people overwhelmingly prefer diagnosticity.

Moreover, when people are allowed to generate their own questions, as opposed to being constrained to select researcher-generated questions, the vast majority of queries cannot logically be classified as confirmatory or disconfirmatory, since they do not favor the hypothesis or its alternatives. Rather, such questions (open-ended and bidirectional) permit wide latitude in the responses of targets. Furthermore, even for the subset of questions that can reasonably be viewed as consistent with the hypothesis, almost none are as constraining or biased as those used in the original research of Snyder and Swann.

Limits of the Two Approaches and Possible Resolutions

Given the conflicting literature, there are problems with endorsing either of the above approaches as the one "most" preferred by social judges. Each is susceptible to criticisms regarding the limitations of their preferred methodologies, and hence the generality of their conclusions. To a certain extent, each program may be susceptible to a confirmatory bias regarding its own preferred explanation.

Difficulties with the confirmatory approach

Several problems with the research procedures of Snyder and Swann (1978) have already been described in the review of the diagnosticity hypothesis. As suggested by Trope and Bassok (Trope and Bassok, 1982; Trope, Bassok, & Alon, 1984), the questions utilized by Snyder and Swann in almost all of their experiments are biased, undiagnostic, and relatively uninformative regarding either the hypothesis or its alternative. Examples of these types of question included "What would you do if you wanted to liven things up at a party?" "What kind of situations do you seek out if you want to meet new people?" "In what situations are you most talkative? What is it about these situations that makes you like to talk?" (p. 1204). The most self-absorbed introvert would have difficulty dispelling the idea that he was somewhat gregarious if a social judge's only data were responses to questions such as these.

Indeed, in the only study from this program of research (Swann & Guliano, 1987, experiment 1) that permitted subjects to generate their own questions, none of the questions generated were as biased as those utilized by Snyder and Swann (1978). Thus, the research paradigm seems to hinge on how subjects utilize questions they would not spontaneously generate themselves.

As discussed before in reviewing Trope and colleague's research, it is also questionable whether the search strategies of hypothesis testers bias the sampled evidence the way Snyder and Swann (1978) suggest. Framing "yes-no" questions so that "yes" responses are evidence for the target's possession of the hypothesized trait does not logically preclude accurate testing of hypotheses, at least not when "no" responses are equally credible evidence against the target's possession of the trait (Fischoff & Beyeth-Marom, 1983; Klayman & Ha, 1987; Trope, Bassok, & Alon, 1984).

To make the above point less abstract, it is useful to consider Wason's study (Wason, 1960), one of the earliest to show a confirmation bias and one which is regularly cited by Snyder (1981; Snyder & Gangstead, 1981; Snyder & Swann, 1978). Wason presented subjects with three numbers (2-4-6) that were related by a rule. The subject's task was to generate further instances to test their best hypothesis about the rule. The experimenter would provide feedback after each set of instances indicating whether it did or did not conform to the rule. When the subject was confident he had determined the rule, he was to tell the experimenter.

Wason found that subjects often failed to determine the rule (increasing numbers) because they tested instances that confirmed but not instances that disconfirmed their own hypotheses. For example, many subjects formulated initial hypotheses like "increasing even numbers," and then proceeded to generate examples that always confirmed the rule (i.e., "8, 10, 12" "100, 102, 104"). Those who did, of course, rarely discovered the correct hypothesis.

Wason's research would appear to suggest that people are prone to using suboptimal methods for testing hypotheses. However, as Klayman and Ha (1987) argue, the utility of _any_ hypothesis-testing strategy is determined by the (often unknowable) relation between instances governed by the hypothesized rule and those covered by the "true" rule and there is no single "best" strategy that guarantees the successful discovery of the true rule for all possible relations between the two rules. Proposing "100, 102, 104" as to test of the hypothesis "increasing even numbers" may be relatively uninformative when the true rule is more general than the hypothesized rule[1], it may be highly informative when the "true" rule is more specific than the hypothesized rule (for example, if the true rule is "increasing sequential even numbers less than 60").

Given our tentative understanding about the nature of social constructs and categories (Holyoak & Gordon, 1984; Kruglanski, 1988; Lingle, Altom, & Medin, 1984) and their relation to social "reality" (Cronbach, 1955; DePaulo, Kenny, Hoover, Webb, & Oliver, 1987; Funder, 1986; Hastie & Rasinski, 1988; Kruglanski, in press), it may far too early to conclude a priori that any one strategy for obtaining social information biases or otherwise distorts social judgments. It is somewhat ironic that Swann's own research (Swann, 1983, 1985; Swann, Griffin, Predmore, & Gaines, 1987) demonstrates the diligence people show in attempting to disconfirm what they perceive to be others' mistaken hypotheses about them. This assumption is central to his theory of preservation of the self-concept (Swann, 1985), yet how such an assumption can be integrated with the assertions of Snyder and Swann (1978) is not clear.

---

[1] By more general I mean that all possible instances governed by the hypothesized rule comprise a subset of all instances governed by the "true" rule.

Beyond questions concerning the logical coherence of the confirmatory formulation, it is possible to question how well Snyder and colleagues data provides support for their assertions. Snyder & Swann (1978) originally speculated that people might use of one of three strategies: confirmatory, disconfirmatory, or equal-opportunity. A confirmatory strategy was defined as one in which people looked only for hypothesis consistent evidence, a disconfirmatory one where people only looked for information inconsistent with the hypothesis, and an equal-opportunity one where people looked for both kinds.

From the findings of each of their studies, Snyder and colleagues concluded that people use confirmatory strategies. However, calculating the number of hypothesis-consistent questions subjects select as a proportion of all questions selected, it becomes apparent that the "confirmatory" hypothesis receives only weak support. Across five studies and 228 subjects, confirmatory questions were selected only a little over 50 percent of the time. Similar proportions of confirmatory questions to total questions are reported by Swann and Guliano (1987, experiment 2). And Swann and Guliano (1987, experiment 1) and Trope, Bassok, and Alon (1984, experiments 1 and 2) found much smaller percentages of confirmatory questions to total questions (note that in these studies subjects were permitted to generate their own questions).

Even restricting our focus to Snyder's experiments (see Table 1), it is not at all clear whether the data best supports the confirmatory hypothesis or the equal-opportunity hypothesis. It can be seen in the table that across all studies, subjects showed only a slight preference for questions consistent with extroversion when testing an extroverted hypothesis, and never showed a preference for introversion-consistent questions when testing the introvert hypothesis.

This overall preference for extravert-consistent questions shown most clearly in Swann and Guliano (1987, experiment 1), where subjects assigned to both introvert and extrovert hypotheses conditions were far more likely to ask questions consistent with the extrovert hypothesis.

The finding that those testing an introvert hypothesis prefer questions consistent with an extravert hypothesis is clearly at odds with predictions generated by the confirmatory formulation. Yet similar findings are obtained in Trope, Bassok, and Alon (1984, experiments 1 and 2). While both Snyder and colleagues and Trope and colleagues have consistently found a bias in favor of extravert-consistent questions, neither has noted it as anything more than an anomaly, when they note it at all. However, within the framework to be presented shortly, I propose that this effect is reflective of a systematic influence on information gathering strategies.

Table 1

Mean Proportion of Hypothesis-Consistent Questions Selected as a function of Assigned Hypothesis:

	Subjects' Hypothesis			
Study	Extrovert		Introvert	
	%	n	%	n
Snyder and Swann (1978)				
Experiment 1	60.8	29	46.5	29
Experiment 2	60.4	20	49.2	20
Experiment 3	54.2	30	na	na
Experiment 4	60.0	15	41.1	15
Snyder and Campbell (1980)	58.6	30	45.8	30
M	58.5	124	46.0	94
Totals	54.1	228		

Difficulties with the diagnostic approach.

The diagnosticity framework appears to enjoy considerable support (Kruglanski & Mayseless, 1988; Skov &

Sherman, 1986). Moreover, since Trope and Bassok (1982, 1984) directly pit diagnosticity and consistency manipulations (Platt, 1964), some conclusions about the role of each appear possible. Across a number of studies diagnosticity accounts for the largest share of systematic variance. Subjects occasionally show slight preferences for questions consistent with features of the hypothesized trait (Skov & Sherman, 1986), but variance accounted for by consistency appears to be substantially less than that accounted for by diagnosticity.

The experiments demonstrating subjects' preference for diagnostic information do suffer from a number of potentially limiting factors however. For example, several of the studies (Bassok & Trope, 1984; Kruglanski & Mayseless, 1988; Skov & Sherman, 1986; Trope & Bassok, 1982) require subjects to choose from lists of questions. Hence, the same limitation present in Snyder and Swann's research is present here; the fact that subjects reliably choose existing questions that are high in diagnosticity does not guarantee that they spontaneously generate such questions. In fact Trope, Bassok, and Alon (1984), who did require subjects to construct question lists, found that across conditions subjects generated questions of relatively moderate levels of diagnosticity. Several of the studies reported above used separate sets of subjects for purposes of generating questions (i.e., Bassok & Trope, 1984;

Kruglanski & Mayseless, 1988) and found that subjects generate questions of both high and low diagnosticity.

A second factor that may potentially limit the generalizability of findings for the primacy of diagnosticity is the emotional investment of the information gatherer with his hypothesis. Subjects in Trope's research typically have little emotional investment in either confirming or falsifying the given hypothesis. Subjects who possess a high interest in either confirmation or disconfirmation may formulate quite different strategies. While it may be true that in many instances social information gatherers do not have a vested interest in the outcomes of their information search, it is also true that in some critical circumstances they can. For example, a teacher who is unsuccessful in teaching a student may be more likely to attend to evidence congruent with the idea that the student is a slow learner than with evidence that the teacher's skills are deficient.

Similarly, an individual motivated to maintain stereotypes may also selectively seek information that maintains the stereotype. Some support for this notion can be found from studies investigating the kinds of information people seek following highly involving success and failure feedback, or following reversible or irreversible decisions for which subjects are personally responsible. In these situations, people show consistent self-serving biases in the kinds of information they attempt seek and consider (Frey, 1978, 1981a, 1981b, 1986; Frey & Rosch, 1984; Pyszczynksi & Greenberg, 1981, 1987; Pyszczynksi, Greenberg, & Holt, 1985; Pyszczynski, Greenberg, & LaPrelle, 1985; Schlenker, 1980).

Pyszczynski and Greenberg (1987) have suggested that different kinds of hypothesis testing processes may characterize individuals with different motivations. The motive for accuracy in decisions, which is presumably best served by diagnostic information search strategies, may not always predominate for the individual:

> Sometimes it is extremely important to arrive at an accurate attribution; other times the need for accuracy is markedly less important or overwhelmed by other more powerful needs and desires. For example, although most jurors who serve in actual trials are undoubtedly highly motivated to arrive at accurate attributions for a defendant's behavior, these same individuals would probably be much less concerned about accuracy when participating in a mock trial. Likewise, the anger and outrage of a relative of a victim of a serious crime may interfere with or overwhelm the juror's desire to draw objective conclusions about the incident in question. (p. 316)

Evidence supporting the notion that the degree to which individuals will search for diagnostic information is contingent on motivation for accuracy comes from Kruglanski and Mayseless (1988, experiment 2). The researchers were interested in whether preference for diagnostic information for testing a hypothesis would vary as a function of one of two motivational sets. "Need for closure" reflects "the striving for clear-cut knowledge on a given topic, and the intolerance of confusion and ambiguity" (p. 9). Because the formulation of alternatives to the given hypothesis is assumed to foster confusion and prevent closure, it was assumed that inducing a need for closure would make subjects less likely to consider alternatives to the hypothesis, and rely more on hypothesis-consistent questions. Kruglanski and Klar (1985) suggest need for closure may originate from individual difference factors such as intolerance of ambiguity, or from situational factors such as lack of time, obligation to reach a decision, or the need for immediate action.

Conversely, a "fear of invalidity" is presumed to be aroused by "the anticipated cost of a mistaken judgment" (Kruglanski & Klar, 1985, p. 42). Fear of invalidity "may make one inclined to generate numerous alternative hypotheses on a given topic and prudently attend to all the relevant evidence one can find" (p. 42). Fear of invalidity is presumed to foster a need for openness, or the need to avoid premature commitment to a particular closure.

Mayseless and Kruglanski (in press, cited in Kruglanski & Mayseless, 1988) found that need for closure inhibited, and need for openness facilitated, the number of hypotheses subjects generated for purposes of explaining a target event. Thus, it was expected that these needs might also be relevant for determining the conditions under which subjects would be more or less likely to prefer diagnostic information.

Judges rated 117 items in terms of how well they discriminated a prototypical architect from a prototypical painter. The items were also rated for their independent associations with a prototypical architect and a prototypical painter. The final list of 32 questions represented orthogonal levels of association with an architect (high versus low), association with a painter (high versus low), and diagnosticity with regard to artists versus painters (high versus low).

Subjects in the fear of invalidity condition were told that ability to identify accurately a person's profession from personality information was an important ability, and that they could receive an extra hour of experimental credit if they could correctly determine the target's profession from his answers to the subject's questions. Subjects in the need for closure conditions were told the experiment was concerned with how quickly people made decisions. A control condition was not given either the accuracy or the speed set of instructions. All subjects tested whether the target was a painter and were told in addition that "the interviewee could, of course, be a member of a different profession; for instance he or she could be an architect."

Results showed that subjects in all conditions tended to prefer hypothesis consistent questions to hypothesis inconsistent ones. More importantly, subjects in the fear of invalidity condition tended to prefer more diagnostic questions than did subjects in the control, who in turn preferred more diagnostic questions than subjects in the need for closure conditions. Thus, preference for diagnostic information for testing a hypothesis did systematically and predictably vary as a function of motivational set.

A final potentially limiting factor from the diagnosticity experiments concerns the extent to which diagnosticity cues present in the experiments are representative of real-world conditions. Trope and colleagues present detailed and simplified bar graphs to subjects in order to permit the assessment of diagnosticity.

This highly artificial manipulation is problematic, because it is not clear what real-world analogs exist for the stimuli they present. Few people in ordinary social life have detailed, clear information concerning the diagnosticity of features available for testing a hypothesis. People probably do prefer diagnostic information, as Trope and Bassok suggest. But under what conditions, if any, do they construct questions that successfully elicit such information remains an unanswered question. It is to just such a question that the present research turns next.

Factors Likely to Affect Strategies of Information Search

What then can be concluded about the strategies of information search that people are likely to employ for testing a hypothesis about a target? Are people guided by a pervasive desire for diagnostic information with regard to the hypothesis and its alternative, as suggested by Trope and Bassok (1982)? Or are people driven to confirm their hypotheses, irrespective of their emotional commitment in the hypothesis or in motivation for accuracy, as suggested by Snyder (1981; Snyder & Gangstead, 1981)? Rather than asking whether people prefer one or another strategy for gathering information, it may be prudent to ask: "What are the conditions under which people are likely to employ these (and potentially other) information gathering strategies.

Schlenker and Weigold (1989) addressed a similar controversy regarding preferences for self-relevant feedback by recommending that the interpersonal activities of people be analyzed within frameworks that account for personal goals, audience factors, and situational constraints and opportunities. Thus, preference for particular kinds of information about self is viewed not as a uniform drive across all situations, but as a variable subject to systematic influences. In the present context of information search strategies, a focus on inputs to the information-search context can help not only to identify conditions under which people will rely on the previously reviewed strategies, but when they will use other strategies as well. For example, when will people gather information inconsistent with a hypothesis? Tentatively, it can be proposed that searches for information inconsistent with the hypothesis will occur when a hypothesis is threatening or undesired, as when people receive undesired feedback about self. This idea has already been alluded to within several frameworks (Frey, 1986; Pyszczynski & Greenberg, 1987; Schlenker, 1984, 1985, 1986a, 1987), although it has only been tested in the social psychological literature for information about self.

Trait Desirability

In a like manner, people may solicit information in a way inconsistent with their hypothesis when the hypothesized state is undesirable or negative. In situations where the target is a loved one, it is easy to imagine motivational forces playing a role in steering inquiry away from an undesired conclusion. However, even in more mundane situations, people may be reluctant to frame questions toward a target in a way that suggests or implies the target indeed possesses the undesired characteristic, particularly if questioners believe equally diagnostic information can be obtained by framing questions to be consistent with an opposite and positively valenced characteristic.

Recall it was noted earlier that information seeking studies using the introvert-extravert trait pair frequently find a bias for questions consistent with introversion, but not extraversion. It is proposed that this is due to a greater perceived social desirability of the characteristic "extraversion" as compared to "introversion." Careful reading of Snyder and Swann's stimulus materials suggest that qualities associated with extraversion are quite desirable, while introvert characteristics appear to be either neutral or somewhat undesirable. To test the notion that trait desirability may influence question consistency, the present research incorporated traits of varying desirabilities as attributes to be tested for in the target. Two of the traits were chosen to be desirable in nature, while two were chosen to be undesirable.

Interpersonal Goal

Consistent with the logic presented above, an important determinant of information seeking should be the actor's goal(s) on the occasion (Schlenker, 1985, 1986a, Schlenker & Weigold, 1989). Previous research generally shares the assumption that subjects attempt to generate diagnostic questions (even Snyder and Swann concede this, although they argue that subjects believe confirmatory questions are diagnostic). However, other goals and motives may often influence the search process (Pyszczynski & Greenberg, 1987) and thereby dictate preferred strategies.

Because information about the hypothesis must typically be obtained from the target himself, self-presentational goals vis a vis the target are likely to be important. A concern for the diagnosticity of the questions asked may be balanced by a concern with appearing likable to the target, especially when the subject expects to meet and/or wishes to make a favorable impression on the target. This is especially likely to be true when the hypothesized state is negative. Concern about the target's impression of the information seeker may also influence the type of questions asked (i.e., open-ended questions like "Tell me about yourself?" versus yes-no questions "Do you like parties?"

Accountability

One final factor that is expected to impact on the social information-seeking process is the personal accountability of the information seeker. Schlenker (1986b; Schlenker & Weigold, in press; Schlenker, Weigold, & Doherty, in press) and Tetlock (1983a, 1983b, 1985; Tetlock & Kim, 1987) have argued for the importance and ubiquity of accountability in social life. While a number of researchers have examined how accountability impacts on the processing of information (Ford & Weldon, 1981; McAllister, Mitchell, & Beach, 1979; Tetlock, 1985; Tetlock & Kim, 1987), there appear to be no studies relevant to whether accountability affects the acquisition of information (Schlenker, 1986b). What research does exist suggests that individuals who are as opposed to are not accountable make greater use of available information (Chaiken, 1980; McAllister et al., 1979; Rozelle & Baxter, 1981; Tetlock, Skitka, & Boettger, in press), even when consideration of available information requires considerable cognitive resources (Tetlock, 1983a; Tetlock & Kim, 1987). Moreover, accountability appears to encourage individuals to work harder at tasks (Doherty, Weigold, & Schlenker, 1989a, 1989b; Schlenker, 1986b) and to be more determined to meet or exceed relevant standards of success (Schlenker, Weigold, & Doherty, in press).

Why do people adjust their actions so when accountable? Being accountable, or answerable, means that one is responsible for fulfilling certain obligations, duties, and other prescriptions for conduct (Schlenker, 1986b). When people are accountable, their actions are subject to evaluations and judgment by audiences. As a consequence, activities that occur in contexts of accountability have greater relevance for an actor's identity, and are thus more likely to impact on the actor's thoughts, feelings, emotions, and behaviors. From the perspective of evaluative audiences, accountability means that actors can be monitored, that their performances can be meaningfully compared to preexisting standards, that a judgment can be made regarding the matching-to-standard of the performance, and that rewards or punishments can then be meted out following the judgment. When people meet or exceed standards in a meritorious fashion, approbation and reward are the norm. Violation of the standard may bring condemnation and punishment.

The pyramid model of accountability (Schlenker, 1986b; Schlenker & Weigold, in press) begins with the proposition that accountability and identity are interrelated. Identity is invariably constructed and maintained in a context of accountability. Self-identifying activities can be regarded as accountings in which one's relevant assets and liabilities are selectively documented and evaluated for the benefit of an audience. When people are accountable for particular events (e.g., job performance, doing one's duty) task performance has implications for identity. When people are not accountable for their actions, events cannot adhere to self, and task performance cannot be used as evidence or documentation supporting or threatening an actors possession of desired attributes, traits, or abilities (Schlenker, 1986b).

More specific to the present study, Schlenker (1986b) proposes that when people are accountable versus not accountable, they will conduct a more intensive search of the environment for information relevant to an anticipated performance. Such a proposition is consistent with both self-identification theory (Schlenker, 1985, 1986a, 1987; Schlenker & Weigold, 1989) and Tetlock's theory of accountability (Tetlock, 1983a, 1983b, 1985, Tetlock & Kim, 1987). Both frameworks suggest that accountability serves to increase task importance, thereby causing people to marshall task-relevant resources. People become more sensitive to the implications of information and analyze it more thoroughly. However, no studies to date have investigated the effect of accountability on information seeking. It can be hypothesized that people will expend greater effort gathering information when they are accountable for a judgment based on the information. This is likely to result in accountable as opposed to unaccountable subjects: (a) gathering more information, (b) gathering better (more diagnostic) information. Some support for these assertions comes from the earlier reported research of Kruglanski and Maseless (1988). Recall that a "fear of invalidity" provoked a more careful information search strategy for subjects who expected to test a hypothesis. Kruglanski and Maseless argue that accountability may be an important factor in producing a fear of invalidity. When people are accountable for their decisions, they reflect on the negative implications of a poor decision, and on the defensibility of their decision processes. Thus, the Kruglanski and Maseless study provides tentative support for the assertion that accountability will impact on information search strategies.

Overview

A study was conducted in order to test several of the preceding assertions. First, to avoid limitations inherent in requiring subjects to select experimenter-constructed questions, subjects were asked to generate their own questions for an interview in order to find out about another person's standing on a trait. There were few limitations on the number or type of questions subjects could ask, except they were to limit questions to a number that could reasonably be answered within half an hour. Each subject drew up a list of potential questions from which he or she choose a final list that, in the experimental conditions, was to be presented to the target. Subjects believed that the target would respond in writing to the questions, and that the subject would administer the questions and collect them from the target.

Three interaction goals were assigned. To manipulate motive to make a favorable impression, some subjects were told to imagine they would be determining standing on a trait for purposes of determining whether or not to share an apartment with the target. They were instructed to ask questions that were as warm and friendly as possible, since the target would be evaluating the subject as a potential roommate as well. Others were told they were determining standing on a trait for purposes of a hiring decision, and that they should be direct. These subjects were specifically told that their only concern should be whether or not the target could be characterized by the trait. Both groups of subjects were told they would give their questions to the target, and would then read over the responses. A control group was told to write questions for a hiring decision, but was told they would not give the questions to a target.

The second factor in the study was accountability. In the experimental conditions, accountable subjects were told they would estimate of the target's standing on the trait after obtaining his answers. Then, the subject would meet with the experimenter to discuss the subject's questions and to receive feedback about the accuracy of his estimate. Nonaccountable subjects were not told they would meet with the experimenter, receive feedback, or talk about their questions. Accountable controls were told they would meet with the experimenter and discuss the questions they selected, but were not told about estimating anyone's standing on a trait.

The final factor was the desirability of the trait being tested. From a list of traits (Alicke, 1985) two highly desirable traits (mature, cooperative) were selected. Bipolar opposites of the desirable traits were used to create undesirable matches (immature, uncooperative). Although much previous work has relied exclusively on the trait pair introvert-extrovert, there is nothing in either the diagnosticity or confirmatory frameworks to suggest that similar processes will not operate for other trait terms as well. Moreover, given the substantial amount of data now available on information search for introvert and extravert hypotheses, it was believed that different trait terms would permit assessment of the generalizability of previous findings. All variables were manipulated between subjects.

CHAPTER 2
METHOD

Subjects

A total of 237 male and female subjects participated in partial fulfillment of a course requirement.

Procedure

During each session groups of four to seven worked at desks separated by partitions. All subjects read informed consent sheets which noted that they would be taking part in an experiment designed to find out how people learn about one another. Instructions explained that the study was investigating how people gather information about one another. Subjects were told "Your task involves generating questions that would be useful for finding out information about another person." Participants would be randomly assigned to play either an interviewer or a respondent (actually every subject received interviewer instructions).

It was noted that people could use a number of ways of finding out about someone, for example

> they may find out about someone by talking to mutual acquaintances. Alternatively, they can withhold judgment about a person and try to observe the individual over different times and in different situations.

However, on some occasions people might have to make quick judgments, as when meeting a stranger at a party or when making a hiring decision. Because the experiment was concerned with such occasions, "you will be asked to gather information about a respondent such that you can make a relatively quick judgment."

Interview Condition

Subjects assigned to the hiring-decision condition were told to imagine that they were a personnel director at a company. The company had decided that a specific trait was an important dimension on which to screen applicants, and they were to determine whether or not the target (an interviewee) possessed the trait. It was emphasized to the subject that his sole objective during the interview was "to learn information about the person that will permit you to accurately judge whether or not the respondent" possessed the trait.

Subjects in the roommate-decision condition were told to imagine that they were interviewing a potential roommate for an off-campus apartment the subject shared with two friends. The friends had agreed that it was important to screen potential roommates on a particular trait. Subjects were told that their primary task was to gather information for making an accurate judgment about the target, but that

> as in real-life situations of this kind, the respondent will also be evaluating you. Therefore, to the extent you can, try to create a positive impression on the respondent. Make your questions as warm and friendly as possible.

Accountability

The degree to which subjects were accountable for their choices was also varied. Accountable subjects were told they would meet with the experimenter following the interview. During the meeting the experimenter would look over and evaluate the subject's question list. The subject would then find out how well his trait judgment matched a personality profile of the target. Subjects who were not accountable did not receive any information about meeting with the experimenter, nor were they told that their questions would looked over or evaluated. Finally, no mention of a personality profile was made.

Trait Desirability

The final manipulation was the valence of the specific trait subjects were to investigate. In the desirable trait condition, subjects were to determine whether or not the respondent possessed one of two highly desirable traits: mature or cooperative. In the undesirable trait condition, subjects were to determine whether the target possessed one of the bipolar opposites of the desirable traits: immature or uncooperative. Desirable traits were obtained from Alicke (1985), who asked judges to rate large numbers of trait terms along dimensions of desirability and controllability. While mature and cooperative were each highly desirable, they varied in the extent to which subjects viewed them as controllable (with cooperative seen as significantly more controllable than mature). The undesirable traits were simply the opposites of the two desirable ones.

After subjects generated questions for the interview, they were given a booklet entitled "Preinterview Questionnaire" (see Appendix A). The booklet contained manipulation checks as well as measures assessing subjects' perceptions of the task, how easy or difficult it was to test for both the assigned trait and the nonassigned traits, how friendly and diagnostic the questions the subject generated were, and base-rate estimates for both the assigned and unassigned traits among Florida psychology students. After completing the questionnaire, subjects were told they would not need to participate in an interview and were debriefed.

Controls

Two control groups were designed to assess the impact of the anticipated interview. The controls were identical to the "hiring decision" experimental groups except for the following. First, controls were told that although their questions would be used in an interview, the subjects themselves would not conduct the interview. Rather, because of time limitations, the questions would be used by other subjects in a later experiment.

Second, because of the obvious inconsistency with control group instructions, accountable controls could not be told that they would evaluate a target's standing on a trait. Rather, these subjects were told only that their questions would be evaluated by the experimenter. Nonaccountable controls were given instructions similar to nonaccountable subjects in the experimental cells. The implications of the potential accountability confound for the analysis are discussed in the next chapter.

Coding

Subjects' protocols were separated into those relevant to either mature-immature, or cooperative-uncooperative. Trained undergraduate judges first classified questions into the following categories (from Bassok & Trope, 1983, and Swann & Guliano, 1987): 1) yes/no questions, or questions designed to be answered with either a yes or no response ("Do you like to go to movies?"), 2) A or B questions, or questions posing a choice between two alternatives ("On a Friday night, would you rather sit at home reading or go to a movie?"), 3) multiple choice questions, or questions providing a range of response options greater than two ("I would rather: a) go to a movie, b) go to the theater, or c) stay home."), and 4) open-ended questions, or questions that do not *a priori* define the subjects response ("What do you like to do on Friday nights?").

Next judges rated each question in terms of its diagnosticity for the relevant trait pair (i.e., distinguishing between a mature versus immature person, or distinguishing between a cooperative versus uncooperative person). Ratings were from 1 (not at all diagnostic) to 5 (extremely diagnostic). If the question was of the nature of a forced choice between pairs (i.e., Yes-No or A/B), judges answered three additional questions. The first question asked for the "probability that anyone would answer 'yes' (true) to the question?" The second elicited the "probability that a mature (or cooperative) person would answer yes" to the question. Finally, judges rated the probability that an immature (or uncooperative) person would answer yes to the question. Judges were blind to protocol conditions except for bipolar trait pair. Within bipolar pairs, judges were blind to the trait for which subjects were testing.

CHAPTER 3
RESULTS

Design and Analyses

A potential confound of the design arose from differences in the manipulation of accountability for experimental and control subjects. This was necessary in order to maintain the beliefs of control subjects that they would not interview another subject. Accountable controls expected to have their questions evaluated, but they did not expect to make a trait judgment. Had the manipulations been equivalent across interview conditions, the design would have been a 2 (desirable versus undesirable trait) x 3 (interview condition: roommate, hiring, control) x 2 (accountable versus not accountable).

In a preliminary analysis the design was treated as a factorial. Instances where accountability achieved statistical significance, either alone or interacting with other variables, were to be reanalyzed with the controls treated as offset cells. However, the initial analysis of all major dependent variables (described in greater detail later) failed to show any significant main effects or interactions for accountability. Therefore, in the analyses that follow, all major dependent variables were analyzed in a 2 (accountable versus not accountable) x 3 (hiring condition versus roommate condition versus control condition) x 2 (desirable versus undesirable trait) factorial design.

Accountability

The manipulation check for accountability assessed responses to the statement "The instructions explicitly told me that I will be meeting with the experimenter to discuss my interview questions." Subjects responses were collected on scales that ranged from 1 false to 9 true. A copy of the manipulation checks is contained in Appendix B. The difference between accountable and unaccountable subjects was significant, $F(2, 226) = 293.73$, $p < .0001$, with the former more likely to judge the statement true than the latter (Ms = 7.8 and 2.8, respectively).

Interview Condition

A main effect of interview condition was highly significant for agreement that "The instructor explicitly told me that my questions should be warm and likable," $F(2, 226) = 135.69$, $p < .0001$. The means for the conditions were as follows: roommate $M = 5.98$, hiring $M = 1.5$, and controls $M = 1.3$. As expected, subjects in the roommate condition differed from subjects in the hiring, $F(1, 160) = 140.94$, $p < .0001$, and control, $F(1, 153) = 144.30$, $p < .0001$, conditions, while the latter groups did not differ, $F(1, 157) < 1$.

Subjects were also questioned about whether they worried "quite a bit about how the respondent would view me." The response scale ranged from 1 _false_ to 9 _true_. The analysis showed a significant main effect of interview condition, $\underline{F}(2, 226) = 3.19$, $\underline{p} < .05$. Simple comparisons suggested that subjects in the roommate condition ($\underline{M} = 3.9$) differed those in the hiring condition ($\underline{M} = 3.0$), $\underline{F}(1, 160) = 18.61$, $\underline{p} < .0001$. Subjects in the control groups, who were asked whether they worried "quite a bit about how the respondent would view someone asking the questions," fell intermediate ($\underline{M} = 3.5$) and did not differ significantly from either experimental group.

Trait Desirability

Trait desirability was assessed by asking subjects to judge the desirability of the four traits used in the experiment. These judgments were subjected to a repeated-measures analysis of variance, with rated trait as a within-subjects factor and accountability, interview condition, and assigned trait (four levels) as between-subjects variables. The ANOVA revealed a significant effect of rated trait, $\underline{F}(3, 639) = 1209.88$, $\underline{p} < .0001$.

A second analysis utilized two repeated measures variables, bipolar pair (cooperative--uncooperative versus mature--immature) and trait desirability (mature--cooperative versus immature--uncooperative). The analysis revealed that mean desirability ratings were significantly different both between bipolar pairs, $F(1, 236) = 8.62$, $p < .004$, and between traits differing in desirability, $F(2, 236) = 1909.37$, $p < .0001$. The two variables did not interact, $F(1, 236) = 1.72$, $p > .19$. Means revealed the mature--immature pair was judged slightly more desirable than the cooperative--uncooperative pair, Ms = 5.1 and 5.0, respectively. And, as expected, the mature--cooperative pair was rated more desirable than the immature--uncooperative pair, Ms = 7.90 and 2.11, respectively.

Other Subject Perceptions

The statement "The instructions specifically told me that I will give my questions to another subject (the respondent) during today's session" showed an expected main effect of interview condition, $F(2, 226) = 155.95$, $p < .0001$. Simple-effects tests showed that the control subjects differed significantly from both the roommate $F(1, 153) = 246.70$, $p < .0001$, and hiring conditions, $F\ (1, 157) = 167.92$, $p < .0001$, while the latter conditions did not differ from one another, $F(1, 160) = 1.66$, $p > .19$. Means for the roommate ($M = 8.0$) and hiring ($M = 7.6$) interview conditions were greater than the mean for the control condition ($M = 2.4$).

Expectancy of making a trait judgement about the target was assessed for responses to the statement "The instructions *explicitly* *told* *me* that my trait judgment will be compared with a respondent's score on a personality scale during today's session." Recall that while accountable subjects in the experimental groups were told that they would make a trait judgment, accountable subjects in the control group, as well as those subjects who were not accountable were not told this. Thus, agreement with the statement was expected to yield main effects for accountability and interview condition, plus an interaction of the two variables.

Table 2

Mean agreement that the Subject would make a Trait Judgment by Accountability and Interview Condition.

	Interview Condition		
Accountability	Roommate	Hiring	Control
Accountable	8.0	6.6	1.5
Not Accountable	2.0	2.3	2.2

Note: Variable measured on a nine-point scale with 1 = false and 9 = true.

These predictions were confirmed (all ps < .0001). As can be seen in Table 2, the simple effect of accountability was highly significant for subjects in the roommate and hiring conditions, $F(1, 75) = 346.30$, $p < .0001$, and $F(1,$

79) = 49.28, $p < .0001$, respectively, but not for those in the control condition $F(1, 71) = 2.27$, $p > .13$. Additionally, accountable controls differed from accountable subjects in the roommate, $F(1, 79) = 357.42$, $p < .0001$, and hiring conditions, $F(1, 78) = 93.94$, $p < .0001$. For unaccountable subjects, however, there were no differences among the three groups (all Fs < 1). There was an unexpected difference between accountable subjects in the roommate and hiring conditions, $F(1, 81) = 5.88$, $p < .02$, with those in the former condition believing it more likely they would make a trait judgment.

Primary Analyses

Given the large number of questions to be scored, the seven student judges did not each evaluate every question. Rather, judges worked independently on separate groups of protocols and made ratings until each subject's questions had been evaluated by at least three judges. The three ratings collapsed across individual judges were used in reliability analyses.

The important dependent variables obtained from judges ratings were: the total number of questions generated by subjects, the types or classifications of questions generated, the diagnosticity of questions generated, the diagnosticity of questions within question types, and the consistency of questions generated with the assigned hypothesis.

Question Type and Number of Questions Generated

The first evaluation was a classification of the subject's question. Judges were provided with the definitions of the four question types and were to determine whether the question was open-ended, multiple choice, A/B, or yes-no. These categories were derived from previous research utilizing subject generated questions (Bassok & Trope, 1983; Swann & Guliano, 1987; Trope, Bassok, & Alon, 1984). Judges were instructed to note questions that did not fit the categories. No judge discovered a question that could not be easily coded into one of the four classifications. Agreement as to the type of question was quite good. For each of the four question types, correlations of the three ratings ranged between .87 and .92.

On average, lists contained just under nine questions ($\underline{M} = 8.7$). The overall number of questions generated was subjected to an analysis of variance using accountability, interview condition, and trait desirability as independent variables. It was expected that subjects a) who were less concerned about the impression they were attempting to make, i.e., those in the control and hiring conditions, as well as accountable versus not accountable subjects, would generate more questions. As can be seen in Table 3, however, these predictions were not confirmed. Instead, the analysis revealed significant main effects of interview condition,

Table 3

Mean Number of Questions Generated by Interview Condition and Trait Desirability

	Interview Condition			
Trait	Roommate	Hiring	Control	M
Desirable	10.2	8.6	9.0	9.2
Undesirable	9.0	7.9	7.7	8.2
M	9.6	8.3	8.4	

F(2, 225) = 3.12, $p < .05$, and trait desirability, F(1, 225) = 4.41, $p < .04$. Simple effects tests within interview conditions revealed that the mean number of questions generated in the roommate condition (M = 9.6) was significantly greater than that for the number generated in either the hiring, M = 8.3, F(1, 160) = 4.90, $p < .03$, or control, M = 8.4, F(1, 160) = 4.24, $p < .04$, conditions. The latter two conditions did not differ. In addition, subjects generated more questions when testing for a desirable versus undesirable trait, Ms = 9.2 and 8.2, respectively.

The mean number of questions generated within question type was analyzed, with question type as a within-subjects variable, and accountability, interview condition, and trait desirability as between-subjects factors. As shown in Table 4, analyses revealed that subjects did not generate equal numbers from each category. Significant main effects of question type, $F(3, 675) = 146.12$, $p < .0001$, and trait desirability, $F(1, 225) = 4.02$, $p < .05$, emerged, along with an interaction of question type and interview condition, $F(6, 675) = 2.35$, $p < .03$.

Profile contrasts showed that subjects asked more open-ended questions ($M = 4.3$) than yes-no questions ($M = 3.14$), $F(1, 236) = 19.48$, $p < .0001$. In turn, yes-no questions were more frequent than a/b questions (a/b question $M = 1.03$), $F(1, 236) = 101.35$, $p < .0001$, and a/b questions were more common than multiple choice questions (multiple-choice question $M = 0.23$), $F(1, 236) = 45.92$, $p < .0001$. Analyses conducted within question types revealed only one significant simple-effect of interview condition, that occurring within yes-no questions, $F(2, 234) = 3.25$, $p < .04$. Further analyses revealed that subjects in the roommate ($M = 3.8$) cell asked significantly more yes-no questions than subjects in the hiring ($M = 2.7$, $F(1, 160) = 5.34$, $p < .03$) or control ($M = 2.9$, $F(1, 152) = 3.84$, $p < .05$) conditions, while the latter two did not differ ($F < 1$).

Table 4

Mean Number of Questions Generated by Question Type and Interview Condition

	Question Type			
Interview Condition	Open Ended	Yes/ No	A / B	Multiple Choice
Roommate	5.0	3.8	0.8	0.1
Hiring	4.0	2.7	1.3	0.3
Control	4.2	2.9	1.0	0.3

Overall Diagnosticity

The next set of analyses were to determine if differences exist in the diagnosticity of the question sets generated by subjects. Prior to analysis of the diagnosticity measure, the ratings of the judges were standardized. This was accomplished by subtracting each judge's mean score from each of his observed scores, and dividing the difference by the standard deviation of the judge's ratings. Next, correlations for each of the three ratings, collapsing across judges, were computed. The correlations were modest, ranging from .47 to .55. A mean of the three ratings was computed for each question, and mean total diagnosticities across all of the subjects' questions were derived and subjected to an analysis of variance (ANOVA).

The expected main effect of interview condition was significant, $\underline{F}(2, 225) = 14.40$, $\underline{p} < .0001$. However the means were not in the predicted directions. It had been hypothesized that subjects instructed to make a good impression (i.e., those in the roommate condition) would be less likely to ask diagnostic questions. Instead, subjects in the roommate condition tended to ask questions that were significantly more diagnostic than the questions generated by subjects from the other two interview conditions. The standardized means for the roommate, hiring, and control conditions were, respectively, .41, -.21, and -.22. No other main effects or interaction terms approached significance.

Simple effects tests showed that the mean diagnosticity of questions in the roommate condition differed significantly from that for subjects in the hiring condition, $\underline{F}(1, 159) = 34.26$, $\underline{p} < .0001$ and the control condition, $\underline{F}(1, 152) = 27.51$, $\underline{p} < .0001$. Means for the control and hiring conditions did not differ, $\underline{F}(1, 157) < 1$. Although it had been expected that accountability would influence the diagnosticity of question set, there were no other main effects or interactions that achieved significance.

Diagnosticity Within Questions Types

The diagnosticity of each question type was calculated by creating a mean of the diagnosticity ratings for all questions of a given type. Because of the limited number of observations for some question types, mean diagnosticity ratings for each type were computed ignoring all manipulated variables. In descending order the diagnosticities of each type were as follows: open-ended questions, $\underline{M} = .21$, yes-no questions, $\underline{M} = -.08$, a/b questions, $\underline{M} = -.33$, and multiple choice questions, $\underline{M} = -.57$. Paired comparison T-tests for question-type diagnosticities revealed that open-ended questions were significantly more diagnostic than both yes-no questions, $\underline{t}(178) = 4.21$, $\underline{p} < .0001$, and a/b questions, $\underline{t}(110) = 2.94$, $\underline{p} < .005$. No other differences were significant.

An analysis incorporating the entire design with question type as a repeated measures factor was impossible because relatively few persons generated questions of all four types (n = 14). Instead, the two most common types of questions, open-ended and yes-no, were subjected to an analysis of variance that included question type as a within-subjects factor and accountability, interview condition, and trait desirability as between-subjects factors. From the original sample, 175 subjects generated at least question of each type.

The analysis revealed a main effect of question type, F(1, 163) = 15.92, p < .0001, confirming that open-ended questions tended to be judged as significantly more diagnostic than yes-no questions. In addition, a main effect of interview condition emerged, F(2, 163) = 11.35, p < .0001 (see Table 4). Of course, the pattern of mean question type diagnosticity within interview condition was similar to that reported earlier for total diagnosticity. Roommate questions were rated as higher in diagnosticity than both hiring questions and control questions. No other main effects or interactions were found.

Table 5

Mean Diagnosticity of Open-ended and Yes-No Questions by Impression Condition.

	Interview Condition		
Question Type	Roommate	Hiring	Control
Open-ended	.63	-.07	.00
Yes-No	.43	-.29	-.23

Consistency of Yes-No questions

By definition, the issue of whether or not a question is "consistent" with a given hypothesis, in the sense suggested by Snyder and Swann (1978), is only logical for yes-no questions. Consistent questions are those for which a "yes" response indicates the presence of attributes predictive of the trait, while inconsistent questions are those for which "yes" indicates attributes predictive of the opposite trait. As noted earlier, yes-no questions were not the most common type of question generated by subjects in the experiment. However, because they did comprise approximately one third of all questions generated, enough data existed to determine whether a significant consistency bias was present. Hence, the subset of yes-no questions was analyzed to determine whether or not they demonstrated consistency with the hypothesized trait, as suggested by Snyder and Swann, or whether some other pattern could be determined.

Judges had been instructed to rate each "yes-no" type question for the probability that someone possessing both the desirable and undesirable trait of the relevant bipolar pair would respond "yes." The dependent variable in this analysis was the difference created by subtracting the judged probability that someone possessing the bipolar opposite of the hypothesized trait would respond "yes" to the question from the probability that a respondent possessing the hypothesized trait would say "yes." Positive numbers indicate consistency with the hypothesized trait (i.e., a "yes" response is more probable if the target possesses the trait), while negative numbers indicate consistency with the alternative trait (i.e., a "yes" is more likely if the target possesses the alternative trait).

Recall that the prediction of the "consistency" framework of Snyder and Swann is that subjects prefer questions that are consistent with their hypotheses. An alternative framework is that subjects will ask questions consistent with a hypothesized trait only when the trait is desirable, and will ask questions inconsistent with the trait when it is undesirable.

The difference scores were subjected to an analysis of variance, again using accountability, trait desirability, and interview conditions as between subjects factors. The analysis provided strong support for the desirability hypothesis. The only significant difference to emerge was a main effect of trait desirability, $\underline{F}(1, 114) = 17.96$, $\underline{p} < .0001$. The mean consistency of yes-no questions for all of the cells of the design are presented in Table 6. For trait desirability, the questions of those testing desirable traits tended to be consistent with the trait ($\underline{M} = 5.9$), while questions of those testing undesirable traits tended to be consistent with the alternative trait ($\underline{M} = -8.7$). Thus, subjects use a confirmatory information-seeking strategy only when testing for a desirable trait. Even for the subset of question that can reasonably be considered for evidence of Snyder and Swann's (1978) consistency bias, there is little support for the notion that people always prefer confirmatory questions.

Table 6

Mean Consistency of Yes-No Questions by Accountability, Trait Desirability, and Impression Condition.

	Interview Condition		
Accountability	Roommate	Hiring	Control
	Desirable Trait		
Accountable	4.9	10.7	2.8
Not Accountable	0.3	-0.4	19.6
	Undesirable Trait		
Accountable	1.1	-14.5	-12.4
Not Accountable	-7.0	-11.5	-5.7

Additional Analyses

Data were obtained for an additional set of measures in order to test alternative explanations for the findings. The additional measures include: (a) trait base-rates in the population from which the expected target was sampled, and (b) the perceived difficulty of testing for the assigned trait.

Trait Base-Rates

Data on base-rates were gathered because the base-rate of a trait within a sampled population may influence the information seeking strategies of subjects. In particular, base-rates may influence the consistency of yes-no questions. Subjects may ask questions consistent with the most likely state of affairs with regard to the target, i.e., ask questions that are consistent with the member of each bipolar pair that has the largest base-rate (Trope, Bassok & Alon, 1984, alluded to this possibility). In turn, manipulations within the design of the study may have influenced subjects' base-rate estimates. Perhaps focusing on a trait because it is contained within the assigned hypothesis affects judgments about the trait's base-rate.

The available data for examining these possibilities were collected as follows. Each subject was asked to estimate the base-rates for his assigned trait and each of the three unassigned traits in a population defined as "introductory psychology students at the University of Florida." This is the population from which subjects in the experiment were (unsystematically) drawn.

The base-rate estimates were then subjected to a repeated measures analysis of variance, with trait base-rate as a within-subjects variable containing four levels, and accountability, interview condition, and assigned trait (four levels) as between subjects variables. The only significant effect was a main effect of trait base-rate, $F(3, 633) = 142.08$, $p < .0001$. Means showed that the base-rate for cooperative was highest, $M = 63.4$, followed by mature ($M = 56.6$), immature ($M = 36.6$), and uncooperative ($M = 17.7$). Profile contrasts revealed significant differences in the base rates for cooperative versus mature, $F(1, 234) = 31.35$, $p < .0001$, mature versus immature, $F(1, 234) = 69.57$, $p < .0001$, and immature versus uncooperative, $F(1, 234) = 40.31$, $p < .0001$.

This analysis revealed two important findings. First, there is little evidence that either the independent variables or testing for a specific trait influenced subjects' estimates of trait base-rates. Second, the significant differences among traits for estimated base-rates suggests an alternative explanation for the previously discussed finding that the consistency of yes-no questions varies as a function of trait desirability. This is because, as the means above make clear, the present research suffers from a confound of trait desirability and trait base-rate. The more desirable traits are associated with larger base-rates. Thus, a "non-motivational" explanation for the phrasing of yes-no questions is that subjects tend to favor questions more consistent with the member of a bipolar pair with the largest base-rate.

One way of disentangling the competing explanations of desirability and base-rate is to determine whether a significant effect of trait desirability remains for yes-no consistency after variance associated with the base-rate estimates for each trait of the relevant bipolar pair is covaried. Two separate analyses were performed, one for each set of subjects assigned within bipolar pairs. Data from subjects assigned to generate questions to test for either cooperative _or_ for uncooperative were analyzed using analysis of covariance (ANCOVA) in which the base rate estimate for cooperative _and_ the estimate for uncooperative were covaried prior to assessing any remaining effects of trait desirability. A similar procedure with the appropriate traits was carried out for subjects testing either member of the mature-immature pair.

The results provide some evidence that trait desirability, and not trait base-rate, is the important influence on consistency. Within the mature-immature pair, neither covariate was significant (base-rate mature, $F(1, 34) = 1.67$, $p > .20$, base-rate immature, $F(1, 34) < 1$), while a significant effect of interview condition remained, $F(1, 34) = 4.14$, $p < .05$. Within the cooperative-uncooperative pair, similar findings were obtained (base-rate cooperative, $F(1, 63) < 1$, base-rate immature, $F(1, 63) < 1$), interview condition, $F(1, 63) = 13.63$, $p < .0005$. Thus, while trait desirability remained an important influence on consistency, base-rates had no perceivable effect.

Expected Difficulty of Testing for the Trait

A second alternative account for consistency findings relates to the difficulty that subjects may have perceived in testing for their assigned trait. Subjects may frame questions in such a way as to be consistent with the member of a bipolar trait pair that is perceived to be easiest to test for. In order to account for this alternative, ratings were obtained from all subjects concerning the difficulty in testing for each of the traits. The measures consisted of subjects' responses for each of the four traits to the question, "In general, how easy or difficult is it to determine whether or not a person possesses each of these traits (by asking them questions about themselves)?" Responses were measured on nine-point scales that ranged from *easy* to *difficult*.

Each of the four measures was then subjected to a repeated measures analysis, with evaluated trait as a within-subjects variable, and accountability, interview condition, and hypothesized trait (4 levels) as between subjects factors. The only significant effect was an interaction of evaluated trait and hypothesized trait, $F(9, 632) = 6.31$, $p < .0001$. Means for the interaction are presented in Table 7. Separate contrasts were performed within hypothesized trait conditions for

Table 7

Mean Difficulty in Testing for a Trait by Evaluated Trait and Hypothesized Trait

	Mean Difficulty Testing for:			
Assigned Trait	Cooperative	Mature	Uncooperative	Immature
Cooperative	**4.42**	3.89	4.49	3.26*
Mature	4.39*	5.32	4.83	4.36*
Uncooperative	4.84*	4.66*	**5.40**	4.48*
Immature	3.89*	5.17	3.98*	**5.48**

Note. Means in the diagonal (in bold) represent mean difficulty in testing for the assigned trait. Means with asterisks differ from the diagonal mean of the same row as tested by orthogonal contrasts, *p* .05.

differences in the mean difficulty of testing for the hypothesized trait versus each of the other traits. Row means in the Table that differ from the mean in the diagonal (the mean rated difficulty for testing for the assigned trait) are indicated with an asterisk. The general pattern of the table is supportive of the notion that for any given trait subjects tended to view their assigned trait as the most difficult to test for. Thus, in contrast with base-rate estimates, difficulty ratings were affected by the manipulations of specific trait.

The next set of analyses were designed to reveal whether differences in estimated difficulty of testing for a trait could account for systematic variance in the consistency of yes-no questions. As with base-rates, subjects were divided into two groups in accordance with bipolar trait pair. Next, the rated difficulty in testing for each member of the bipolar pair was covaried in order to determine whether: 1) the covariate accounted for a significant amount of variance in the consistency variable, and 2) whether the covariate reduced the significance of trait desirability on consistency.

With respect to the first question, neither covariate within either condition achieved significance (Fs < 1). With respect to the second question, the effects of trait desirability remained significant in the cooperative-uncooperative conditions, $F(1, 72) = 15.28$, $p < .0002$, and remained marginally significant in the mature-immature conditions, $F(1, 46) = 2.95$, $p < .09$. Thus, the difficulty in testing for a given trait does not appear to be a rival explanation for findings related to trait desirability and consistency.

CHAPTER IV
DISCUSSION

In evaluating the implications of the present study and potential future directions, it may be useful to summarize the original predictions for the study and compare them to the findings. Trait desirability was expected to influence the consistency of yes-no questions; specifically, it was believed that questions would be consistent only when the hypothesized trait was desirable. The findings supported this hypothesis quite strongly. Not only were subjects only consistent when the trait was desirable, they were strongly inconsistent when the trait was undesirable. Accountability was expected to influence the quality and quantity of questions that subjects generated. No support was found for these expected relationships. Indeed, there was little evidence that the accountability manipulation influenced the question lists of subjects on any of the dependent measures. Interpersonal goal as manipulated by interview condition was expected to influence the diagnosticity of the questions subjects choose to ask. This happened, but the specific condition under which questions were _most_ diagnostic was one it was expected would be _least_ diagnostic. In the following sections, each of these issues is dealt with in greater detail. The concluding section points to future directions for research in accountability and information search strategies.

What was Thought Would Happen and Did

A review of previous findings showing differential consistency patterns for introversion as compared to extraversion suggested that subjects might show consistency for a hypothesized trait only when the trait was desirable. At least two findings argue against consistency as a primary "motive" in information gathering strategies, and thus pose problems for the consistency framework of Snyder, Swann, and colleagues. First, subjects generated relatively few yes-no questions, which are the only type that could reasonably be evaluated for consistency. Instead, they most frequently choose to ask open-ended questions, which are the least constraining type with regard to target responses. Perhaps because they are so unconstraining, judges evaluated open-ended questions quite favorably in their diagnosticity ratings. Open-ended questions were seen as significantly more diagnostic than the next two most common categories of questions, yes-no and a/b.

Second, and perhaps more problematic for the consistency framework, is the fact that analyses of the subset of questions that do lend themselves to consistency assessment show that subjects prefer consistent yes-no questions only when their hypothesized trait is desirable.

When the trait is undesirable, subjects strongly prefer inconsistent questions. This support for the trait desirability hypothesis helps explain why previous research has found consistency effects to be relatively strong for the trait extraversion, but weak or absent for the trait introversion. Although empirical proof is lacking, there is reason to believe that many subjects consider extraversion to be a more desirable characteristic than introversion. Certainly such an assumption can be easily tested by obtaining ratings of the desirability of each of these two traits.

The present data also argue against two alternative explanations for consistency findings, i.e., trait base-rates, and difficulty in testing for a trait. Trait base-rates, but not trait difficulty estimates, were found to be correlated with desirability. However, covariance analyses revealed that base-rates and trait difficulty ratings did not account for systematic variance in question consistency, and did not attenuate the impact of trait desirability. Caution is warranted, since the covariance analyses do not offer conclusive proof that base-rates and trait difficulty are unimportant. However, Snyder and Swann (1978, experiment 3) experimentally manipulated base-rates and found they were not significant predictors of question choice for the traits of introversion and extraversion.

Why do people frame yes-no questions to be consistent with the most desirable trait of a bipolar pair? Goffman (1967) suggested that interactions are governed by a "rule of considerateness" whereby the person tends to "conduct himself during an encounter so as to maintain both his own face and the face of the other participants" (p. 11). Goffman's analysis stressed the importance interactants place on maintaining the face not only of self but of others in the interaction. Similarly, Homans argued that "the exchange of rewards takes a larger share in social behavior than the exchange of punishments, if only because the latter, when at all damaging, puts an end to social behavior" (1959, p. 57).

Interaction participants may consider the form of a question, i.e., its consistency with a given trait, to be indicative of the questioner's "best guess" about the target's attributes. Thus, a question phrased to be consistent with an undesirable trait may suggest the interviewer considers it more likely that the respondent possesses that trait than its bipolar opposite, and such phrasing is likely to be viewed as insulting. When people have little reason to do otherwise (i.e., when they don't wish to deliberately insult the target, or when they don't want to create feelings of unease in the target), it can be expected that they will frame questions in a way that implies the most desirable state of affairs about the target.

What Happened but Wasn't Expected

The finding that interview condition significantly affected question diagnosticity confirmed an initial assumption of the research, namely that interpersonal goals would affect the quality of questions subjects would generate. However, it was hypothesized that instructions to make a good impression would reduce the diagnosticity of subjects' question sets. Instead, questions were most diagnostic when subjects were motivated to make a good impression. Indeed, subjects in the roommate condition not only asked better questions, they asked more questions, and they were more likely to generate the types of questions that maximized response diagnosticity. Why?

It is difficult to interpret these findings because the interview condition manipulation suffered from a confound. Subjects in both the hiring and control conditions, but not in the roommate conditions, received instructions relevant to a hiring decision. The results may therefore have obtained because of something related to what was intentionally manipulated (i.e., self-presentation concerns) or because of something related to a nuisance variable (the hiring versus roommate "sets"). In retrospect, this confound could have been handled in one of two ways.

The first would have been to standardize the "set" for the entire experiment (i.e., have all subjects make a hiring decision, and still manipulate impression concerns). This possibility was considered at the initial stages of the study and rejected for several reasons. First, for whichever set was selected, one of the impression manipulations would have been highly artificial given a prototypical "script" for the set. For example, job interviews are typically not circumstances under which the interviewer tries hard to make a good impression, while roommate interviews are typically not circumstances under which a target is grilled with little concern for personal feelings. Clearly, these circumstances are not precluded, and may even be somewhat common. However, it was felt that these did not fit with the prototypical scripts for the situations.

If the impression manipulation did appear somewhat more artificial under one interview condition than another, then the manipulation would still have suffered from a confound, namely, impression concern with the fit of the scenario to the prototype. Other sets besides a hiring decision and a roommate decision could have been used, but the hiring and roommate sets appeared to be quite representative of common social situations in which people actively gather information relevant to a hypothesis.

The two sets were also chosen in order to provide as powerful a manipulation of interpersonal goal as possible. Since this research is the first to demonstrate systematic variance in question diagnosticity as a function of interpersonal impression motives, concern with the impact of the manipulation appeared as warranted as concern with the precision with which conclusions about the manipulation can be drawn. Of course, now that the utility of the manipulation for affecting diagnosticity has been established, it becomes vitally important to determine the specific processes involved.

A second, better way of dealing with the confound would have been to include an additional set of controls who did not expect to give their questions to a target, but who were given roommate instructions. This would have handled each of the concerns raised above and still permitted the ability to disentangle the effects of "set" and interpersonal goal. This option was also considered, but would have required collecting data from several dozen more subjects and coding of several hundred more questions.

Thus, the present study cannot provide definitive answers about the role of interpersonal motives on information search strategies, since the findings related to interview condition raise far more questions than they answer. Yet it does show that these are important. Follow-up studies will be necessary to provide clearer explanations for what has been uncovered. Nevertheless, it is possible to speculate. Explanations can be broken down into those relevant to the desired manipulation of interpersonal goal, and those relevant to the nuisance variable "set."

Explanations Attributable to Impression Concerns

Some explanations for the effect of interview condition are relevant to impression concerns. First, subjects charged with being unthreatening and friendly may have found it easier to think of questions because they presumably had less to worry about in the upcoming interaction. Hiring condition subjects might have worried about looking obnoxious to a peer, creating anxiety and reducing the quality of performance on the question generation task (Baumeister, 1984; Baumeister & Steinhilder, 1984). This line of reasoning, however, is undermined by the advantage of subjects in the roommate condition over the controls, who did not have to worry about the interview at all. Presumably, control subjects experienced the least amount of anxiety while generating questions.

Second, the questions of subjects wishing to make a good impression may have been less easily deciphered as to the hypothesized trait, leaving less possibility for demand characteristics. Thus, even if the questions were equally unsettling or threatening, those generated in the roommate condition may have been less likely to give away the hypothesis. This, in the eyes of the judges, may have diminished the likelihood that responses would be carefully monitored for their "correctness" with respect to the hypothesized trait.

Explanations Attributable to "Set"

Other explanations for the superior quality of the questions in the roommate cells may have been relevant to the unique instruction set. The context of finding out the attributes of a potential roommate may have been more captivating for subjects than finding out about a potential job hire. Subjects may have had more personal experience with roommate as opposed to hiring decisions. This personal experience may have made the roommate scenario easier to imagine, or may have raised the subject's expectations of being able to successfully execute the task. Even in circumstances where subjects had experience with both decisions, they may view making a decision about a living partner as a more important choice than making a decision about an employee. Any of these factors, alone or in combination, may have produced the observed differences. Further research is necessary to determine the role that each of these possibilities may have played.

What was Thought Would Happen but Didn't

The effects of accountability in the present research did not achieve significance for any of the major dependent variables, either alone or in interactions with the other manipulations. Why? Several possibilities exist. The first two speculations focus on relatively unique factors in the experiment that may have reduced the impact of the accountability manipulation, while the last focuses on a more general relation between accountability and intrinsically rewarding tasks.

First, the anticipated interaction with the target may have seriously weakened the impact of the manipulation. Previous research (Schlenker, 1986b; Weigold & Schlenker, 1989) has shown that accountability makes individuals view tasks as more important by making salient the implications of task performance for identity. In the present research, however, the subject's most immediate concern vis-a-vis task performance may not have been his expected interaction with the experimenter, to take place at the conclusion of the experiment, but rather his anticipated interaction with the target, which was far more immediate.

Beyond the immediacy of the subject-target interaction, it is possible that subjects believed they stood to gain and/or lose more as a result of their peer interaction. After all, they might expect to see the target outside of the experiment, perhaps in class. The subject also might expect to themselves with a target they would be highly motivated to impress. In comparison, the experimenter may have been viewed as a less important audience. A potential problem with the above line of reasoning, however, arises from the data of controls who did not anticipate interactions with a target. If accountability had been a more impactful manipulation for these subjects, the results would have contained significant interview condition x accountability interactions whereby the differences between accountable and nonaccountable subjects were minimal in the experimental conditions but powerful in the control conditions. This did not happen.

A second explanation for the failure of accountability may rest in the nature of the task. Previous research that has found significant effects of accountability on performance (Doherty, Weigold, & Schlenker, 1989a, 1989b; McAllister, Mitchell, & Beach, 1972; Tetlock & Boettger, in press; Tetlock, Skitka, & Boettger, in press) has typically either implicitly or explicitly drawn attention to the significance of performance for important identity images. Thus, failure on a task in the Doherty et al. paradigm implied lack of an important attribute, while success implied possession of the attribute. In the present research, no attempt was made to link performance on the task or on the subsequent judgment to important images. Thus, accountable subjects in the present experiment may have been insufficiently focused or attentive to an implication of success or failure for identity.

Third, the failure of the accountability manipulation to affect the question lists may be due to the fact that the task itself was perceived as involving and interesting. It has been suggested that when people are internally motivated to do well on a task, or when a task is intrinsically interesting (Amabile, 1985, Doherty, Schlenker, & Weigold, 1989a, 1989b; Schlenker, 1986b; Schlenker & Weigold, in press), accountability will have less of an impact than when a task is uninvolving. This may happen for a number of reasons. Intrinsically motivating tasks may reduce self-awareness, drawing attention away from self and the implications of performance for self (Deci & Ryan, 1980; Wicklund, 1975). Conversely, tasks on which people are internally motivated to do well may increase accountability to self, thereby making salient personal standards for task outcomes (Carver & Scheier, 1985; Greenwald & Breckler, 1985; Schlenker, 1982, 1985, 1986a; Schlenker, Weigold, & Doherty, in press a).

In sum, a better test of the effect of accountability on information seeking may involve reducing the impact of extraneous social pressures on subjects, and may require making more explicit the relation between performance on a task and relevant public and private components of identity.

Information Search Strategies: Future Directions

The findings of the study suggest the importance of multiple methodologies for assessing interpersonal processes. It now appears likely that the conclusions drawn from the research program of Snyder and Swann are of questionable generalizability (Skov & Sherman, 1986; Trope, Bassok, & Alon, 1984). When subjects are allowed to generate questions for purposes of testing a hypothesis, the results differ substantially from those reported by Snyder & Swann (1978).

Equally suspect are claims that people are exclusively motivated by diagnosticity concerns. By varying interpersonal goal, the present study found systematic differences in the quantity and quality of questions generated by subjects. While Trope (1986) has acknowledged that subjects may gather information under motivations other than a need for accuracy, empirical support is only beginning to accumulate (Kruglanski & Mayseless, 1988; Pyszczynski & Greenberg, 1981; Tetlock, 1983b, 1985). Nonetheless, a growing number of theoretical orientations strongly suggest that a desire for accuracy may often not be a primary goal in information seeking (see Pyszczynksi & Greenberg, 1987) or that lay definitions of "accuracy" must account not only for what is believable, but what serves the goals, needs, and wants of the actor (Schlenker, 1986a; Schlenker & Weigold, 1989).

The present experiment is hardly exempt from serious limitations. In addition to those already discussed regarding the confounds within the interview condition manipulation, the methodology of the experiment may have seriously distorted consistency findings. For example, findings obtained from preinterview lists of questions probably have greater relevance to structured interview settings (those following a preconstructed outline) rather than free-form conversation. Hypothesis testing during conversation may be interactive, in the sense that responses to early questions strongly influence the nature of subsequent questions. Thus, a first response that supports a given state of affairs may prompt a series of questions that are consistent with that state of affairs. Future research would do well to test results obtained here in the context of "on-line" interactions.

At a more general level, the concept of diagnosticity raises several questions. For example, by what standard are questions "diagnostic" of traits? Are diagnostic questions ones that would accurately classify people in terms of their distributions on personality scales? Such a reference point assumes that trait meanings are consistent across people, and across people and personality scales (i.e., that people and scale constructors are attempting to measure the same things). It also assumes the validity of the scales.

Diagnosticity in the present experiment was operationalized with reference to judge's ratings, but such judgments measure diagnosticity only in a subjective sense. After all, why shouldn't biases characterizing the responses of subjects be equally likely to manifest in judges. In point of fact, diagnosticity in the present study (and in others of this type) is analogous to face validity. Perhaps a way around the issue of trait diagnosticity is to evaluate the predictive ability of questions designed to predict behaviors. In this circumstance, objective measures of diagnositicity could be obtained (Funder, 1986; DePaulo & Kenny. 1987), i.e. the predictive validity of questions with respect to the behaviors in question.

Above and beyond difficulties associated with diagnosticity, we know relatively little about how well people can assess information obtained during "on-line" interactions. What information exists suggests interactions involve considerable cognitive "effort" and place heavy demands on attention and memory (Baumeister, Hutton, & Tice, in press; Gilbert, Pelham, & Krum, 1988). Hence, data obtained through conversation are likely to affect judgments differently than data obtained through written communications. For example, the former would be more susceptible to memory biases such as recency and primacy effects (Tetlock, 1983b) and biases associated with interpersonal attraction (Bersheid & Walster, 1977) or attribution errors (Ross, 1977). Conversational as opposed to transcribed interactions may direct attention toward relatively uninformative channels of communication, a phenomenon demonstrated in lie-detection research (DePaulo, Stone, & Lassiter, 1985).

While many questions obviously remain, it is useful to return to issues raised in the opening chapter to assess the implications of the study. First, no evidence was found that the strategies people employ in information search create biases with regard to evaluating a hypothesis. In particular, there was nothing in the data to support the assertion of Snyder and Swann (1978) that information search strategies promote negative stereotypes. Even if it could be assumed (and it probably cannot) that biases in yes-no questions constrain the answers of targets, the results of the present research would argue for a "positivity" bias rather than a consistency bias. That is, people are more likely to ask questions consistent with a desirable versus undesirable trait.

In addressing the question of the kinds of information people most prefer when engaged in search strategies, the answer is less clear. However, because systematic differences in question type, number of questions, and question diagnosticity across conditions were observed, it may be useful to reframe the question as "What kinds of information under what conditions do people most prefer?"

Issues related to information seeking under conditions of high memory load were not addressed in the present study, since subjects who expected to receive responses believed they would be obtained in written format. However, the question does suggest the utility of comparing strategies observed in the present study with conversational information-seeking strategies.

Finally, the study argues strongly for making cautious generalizations from findings in non-social research to social contexts. Subjects faced with testing a hypothesis about another person, especially when such information must be obtained from the target, must incorporate strategic self-presentation concerns along with needs for accuracy. The unique requirements of interpersonal interaction therefore appear to be at least as important in determining search strategies as nonsocial factors.

APPENDIX A
MANIPULATIONS AND INSTRUCTIONS

General Instructions Given to Subjects in the Hiring Conditions

You have been assigned the role of QUESTIONER. Others have been given the role of RESPONDENT. The decision as to who would play these roles was made randomly.

Your task involves finding out information in order to make a judgment about another person. We will explain more about this shortly.

It is often important to know whether or not others possess certain traits and abilities. For example, hiring decisions require relatively rapid judgments concerning whether an applicant possesses desired job-relevant characteristics.

Given motivation to find out about someone else, people may use many different strategies for gathering information. For example, they may find out about someone by talking to mutual acquaintances. Alternatively, they can withhold judgment about a person and try to observe the individual over different times and in different situations.

Many occasions, however, demand that people make quick judgments about others. A typical instance is when someone is deciding whether to hire a job applicant. In this experiment, you will be asked to gather information about a respondent such that you can make a relatively quick judgment. Please turn the page to find out more about the specific task.

Summary Given to Accountable Subjects

1. Generate questions that you believe will allow you to accurately judge the extent to which the respondent possesses the trait.

2. Record the questions in the order you will ask them on the paper(s) marked "Interview Questions." Leave enough room between questions for the respondent's answers.

3. Answer the preinterview rating form.

4. Go to the interview room. The experimenter will bring you and a respondent to an interview chamber. You and the respondent will have approximately 15-20 minutes for the interview.

5. Return to the lab. You will be given sheets on which to record your trait judgment.

6. Discuss your information gathering strategy with an experimenter. Learn how well your judgment matches the personality score from the subject's pretest.

Summary Given to Nonaccountable Subjects

1. Generate questions that you believe will allow you to accurately judge the extent to which the respondent possesses the trait.

2. Record the questions in the order you will ask them on the paper(s) marked "Interview Questions." Leave enough room between questions for the respondent's answers.

3. Answer the preinterview rating form.

4. Go to the interview room. The experimenter will bring you and a respondent to an interview chamber. You and the respondent will have approximately 15-20 minutes for the interview.

5. Return to the lab. You will be given sheets on which to record your trait judgment.

Instructions Given to Accountable Subjects Testing for Mature in the Hiring Condition

Imagine that you are a personnel director for a corporation. Research has determined that maturity is important dimension on which to screen employees. Therefore, the company would like you to assess the degree to which a person you will interview is MATURE. Your sole objective in the interview is to obtain information about the person that best permits you to accurately judge whether or not the respondent is MATURE.

> MATURE (def): A mature individual is someone who can be counted on to handle important duties without constant monitoring or pressure. Mature individuals take responsibility for their mistakes, and do not try to pawn off failures on other people or bad luck. In addition, they try to learn from their errors rather than make excuses for them.

Although there is no minimum or maximum number of questions you are constrained to ask, you should strive to a) ask as many as necessary to form a reasonably accurate judgment, b) not ask so many that the interview would take an inordinate amount of time (20 minutes or so).

Following the interview, you will judge the respondent's maturity. Your judgment WILL NOT be seen by the respondent. However, it will be compared to the respondent's score on the Maturity Factor of the Griggs, Sloan, & Samson Inventory (c), a widely used and well-validated personality measure. The experimenter will tell you how well your judgment matches the objective score, and will ask you about the questions you generated for the interview. Be prepared to discuss your questions and your judgments at this time.

Consider the questions you wish to ask carefully. Once you have composed this list, you will have to stick to it during the interview (just as in real-life personnel directors usually ask the same questions of all job applicants). You will not be allowed to ask questions that you have not prepared in advance. You may wish to jot down a list of questions on the legal pad before selecting the final set for the interview. Be sure to put the final set that you use on the white paper marked "INTERVIEW QUESTIONS." In approximately 30 minutes, the experimenter will bring you and a respondent to a room to begin the interview.

Instructions Given to Nonaccountable Subjects Testing for Immature in the Hiring Condition

In the present experiment, imagine that you are a personnel director for a corporation. Research has determined that immaturity is important dimension on which to screen employees. Therefore, the company would like you to assess the degree to which a person you will interview is IMMATURE. Your sole objective in the interview is to obtain information about the person that best permits you to accurately judge whether or not the respondent is IMMATURE.

> IMMATURE (def): An immature individual is someone who cannot be counted on to handle important duties without constant monitoring or pressure. Immature individuals usually will not take responsibility for their mistakes, and often try to pawn off failures on other people or bad luck. In addition, they rarely learn from their errors since they constantly make excuses for them.

Although there is no minimum or maximum number of questions you are constrained to ask, you should strive to a) ask as many as necessary to form a reasonably accurate judgment, b) not ask so many that the interview would take an inordinate amount of time (20 minutes or so).

Following the interview, you will judge the respondent's immaturity. Your judgment WILL NOT be seen by the respondent.

Consider the questions you wish to ask carefully. Once you have composed this list, you will have to stick to it during the interview (just as in real-life personnel directors usually ask the same questions of all job applicants). You will not be allowed to ask questions that you have not prepared in advance. You may wish to jot down a list of questions on the legal pad before selecting the final set for the interview. Be sure to put the final set that you use on the white paper marked "INTERVIEW QUESTIONS." In approximately 30 minutes, the experimenter will bring you and a respondent to a room to begin the interview.

General Instructions Given to Subjects in the Roommate Conditions

You have been assigned the role of QUESTIONER. Others have been given the role of RESPONDENT. The decision as to who would play these roles was made randomly.

Your task involves finding out information in order to make a judgment about another person. We will explain more about this shortly.

It is often important to know whether or not others possess certain traits and abilities. For example, smooth interactions on a blind date require relatively rapid judgments as to whether a stranger possesses compatible qualities and interests.

Given motivation to find out about someone else, people may use many different strategies for gathering information. For example, they may find out about someone by talking to mutual acquaintances. Alternatively, they can withhold judgment about a person and try to observe the individual over different times and in different situations.

Many occasions, however, demand that people make quick judgments about others. A typical instance is when someone is introduced to a stranger at a college orientation session or at a party. In this experiment, you will be asked to gather information about a respondent such that you can make a relatively quick judgment. Please turn the page to find out more about the specific task.

Instructions Given to Accountable Subjects

Testing for Cooperative in the Roommate Condition

Imagine that you and two friends are looking for a fourth roommate to share an off-campus apartment. Because you and your friends have agreed it to be an important dimension on which to screen roommates, you are trying to assess the degree to which the respondent is COOPERATIVE.

A major part of your task is to make an accurate judgment as to whether or not they are UNCOOPERATIVE. However, as in real-life situations of this kind, the respondent will also be evaluating you. Therefore, just as when you are meeting someone in real life, try to put your best foot forward and be as likable as possible during the interview.

> COOPERATIVE (def): A cooperative person gets along well with others. These individuals tend to think in terms of the group rather than in terms of themselves (i.e., how can I help our group do well, versus how can I make myself look good). Cooperative people tend to enjoy group situations since they don't mind sharing the credit when things go well, and they are not constantly concerned with doing things their own way.

Although there is no minimum or maximum number of questions you are constrained to ask, you should strive to a) ask as many as necessary to form a reasonably accurate judgment, b) not ask so many that the interview would take an inordinate amount of time (20 minutes or so).

Following the interview, you will judge the respondent's cooperativeness. Your judgment WILL NOT be seen by the respondent. However, it will be compared to the respondent's score on the Cooperativeness Factor of the Griggs, Sloan, & Samson Inventory (c), a widely used and well-validated personality measure. The experimenter will tell you how well your judgment matches the objective score, and will ask you about the questions you generated for the interview. Be prepared to discuss your questions and your judgments at this time.

Consider the questions you wish to ask carefully. Once you have composed this list, you will have to stick to it during the interview (just as in real-life interviewers usually ask the same questions of all respondents). You will not be allowed to ask questions that you have not prepared in advance. You may wish to jot down a list of questions on the legal pad before selecting the final set for the interview. Be sure to put the final set that you use on the white paper marked "INTERVIEW QUESTIONS." In approximately 30 minutes, the experimenter will bring you and a respondent to a room to begin the interview.

Instructions Given to Nonaccountable Subjects Testing for Uncooperative in the Roommate Condition

Imagine that you and two friends are looking for a fourth roommate to share an off-campus apartment. Because you and your friends have agreed it is an important dimension on which to screen roommates, you are trying to assess the degree to which the respondent is UNCOOPERATIVE.

A major part of your task is to make an accurate judgment as to whether or not they are UNCOOPERATIVE. However, as in real-life situations of this kind, the respondent will also be evaluating you. Therefore, just as when you are meeting someone in real life, try to put your best foot forward and be as likable as possible during the interview.

> UNCOOPERATIVE (def): An uncooperative person rarely gets along well with others. These individuals tend to think in terms of themselves rather than in terms of the group (i.e., how can I help myself to look good, versus how can I help the group to do well). Uncooperative people tend to avoid group situations since they don't like sharing the credit when things go well, and they are unusually concerned with doing things their own way.

Although there is no minimum or maximum number of questions you are constrained to ask, you should strive to a) ask as many as necessary to form a reasonably accurate judgment, b) not ask so many that the interview would take an inordinate amount of time (20 minutes or so).

Following the interview, you will judge the respondent's cooperativeness. Your judgment WILL NOT be seen by the respondent.

Consider the questions you wish to ask carefully. Once you have composed this list, you will have to stick to it during the interview (just as in real-life interviewers usually ask the same questions of all respondents). You will not be allowed to ask questions that you have not prepared in advance. You may wish to jot down a list of questions on the legal pad before selecting the final set for the interview. Be sure to put the final set that you use on the white paper marked "INTERVIEW QUESTIONS." In approximately 30 minutes, the experimenter will bring you and a respondent to a room to begin the interview.

APPENDIX B
DEPENDENT VARIABLES

Please help us by answering the following questions as best you can. In a moment, we will begin the interview.

1. How would you rate the experiment so far:

CHALLENGING | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | EASY

THREATENING | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | SAFE

INVOLVING | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | UNINVOLVING

CONFUSING | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | CLEAR

2. The questions I have generated for this interview are:

INDIRECT | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | DIRECT

CREATIVE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | UNCREATIVE

USEFUL | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | USELESS

3. In generating the questions, I:

a. Relied heavily on the definition of the trait supplied by the experimenter:

FALSE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | TRUE

b. Relied heavily on my own definition of the trait

FALSE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | TRUE

c. Worried quite a bit about how the respondent would view me:

FALSE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | TRUE

d. Worried quite a bit about how the experimenter would view me:

FALSE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | TRUE

e. Was very concerned I would not think of enough questions to ask:

FALSE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | TRUE

f. Was very concerned I would write down too many questions:

FALSE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | TRUE

g. Was concerned I would insult the respondent:

FALSE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | TRUE

4. The respondent's answers to my questions are likely to be:

INFORMATIVE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | UNINFORMATIVE

5. The respondents know that you are trying to find out about a trait, but do not know which one. Given the questions you have generated, how likely is it that after the interview the respondent will guess you were trying to find out about his or her degree of:

MATURITY
LIKELY | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | UNLIKELY

IMMATURITY
LIKELY | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | UNLIKELY

COOPERATIVENESS
LIKELY | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | UNLIKELY

UNCOOPERATIVENESS
LIKELY | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | UNLIKELY

6. Given the questions that you will ask, the respondent is likely to consider you:

FRIENDLY | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | UNFRIENDLY

UNKIND | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | KIND

SENSITIVE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | INSENSITIVE

7. In general, how easy or difficult is it to determine whether or not a person possesses each of these traits (by asking them questions about themselves)?

MATURITY
EASY | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | DIFFICULT

IMMATURITY

EASY | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | DIFFICULT

COOPERATIVENESS

EASY | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | DIFFICULT

UNCOOPERATIVENESS

EASY | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | DIFFICULT

8. How important is it for you to do well when you make your trait judgment?

IMPORTANT | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | UNIMPORTANT

9. If you pause for a moment to think about the interview, you feel:

CALM | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | NERVOUS

HAPPY | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | SAD

COLD | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | WARM

STRONG | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | WEAK

GOOD | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | BAD

AUTHENTIC | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | FALSE

DISHONEST | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | HONEST

CONTROLLED | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | CONTROLLING

POWERLESS | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | POWERFUL

10 In most social situations, which of the following opposites best characterizes you:

EXTRAVERT | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | INTROVERT

HUMBLE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | ARROGANT

ANXIOUS | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | AT EASE

11. Following the interview, the respondent will probably view me as:

STRONG | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | WEAK

GOOD | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | BAD

CALM | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | NERVOUS

HAPPY	1	2	3	4	5	6	7	8	9	SAD
COLD	1	2	3	4	5	6	7	8	9	WARM
AUTHENTIC	1	2	3	4	5	6	7	8	9	FALSE
DISHONEST	1	2	3	4	5	6	7	8	9	HONEST
CONTROLLED	1	2	3	4	5	6	7	8	9	IN CONTROL
POWERLESS	1	2	3	4	5	6	7	8	9	POWERFUL

12. The task of finding out about the characteristics of someone else is:

EASY	1	2	3	4	5	6	7	8	9	DIFFICULT
ENJOYABLE	1	2	3	4	5	6	7	8	9	UNENJOYABLE
REALISTIC	1	2	3	4	5	6	7	8	9	UNREALISTIC
COMPLEX	1	2	3	4	5	6	7	8	9	SIMPLE
DEMANDING	1	2	3	4	5	6	7	8	9	EFFORTLESS

13. Most college students are:

QUITE MATURE	1	2	3	4	5	6	7	8	9	NOT AT ALL MATURE
QUITE COOPERATIVE	1	2	3	4	5	6	7	8	9	NOT AT ALL COOPERATIVE
QUITE UNCOOPERATIVE	1	2	3	4	5	6	7	8	9	NOT AT ALL UNCOOPERATIVE
QUITE IMMATURE	1	2	3	4	5	6	7	8	9	NOT AT ALL IMMATURE

14. People who know me well would probably judge me to be:

QUITE MATURE	1	2	3	4	5	6	7	8	9	NOT AT ALL MATURE
QUITE COOPERATIVE	1	2	3	4	5	6	7	8	9	NOT AT ALL COOPERATIVE
QUITE UNCOOPERATIVE	1	2	3	4	5	6	7	8	9	NOT AT ALL UNCOOPERATIVE
QUITE IMMATURE	1	2	3	4	5	6	7	8	9	NOT AT ALL IMMATURE

15. How desirable are each of the following traits:

COOPERATIVE
UNDESIRABLE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | DESIRABLE

UNCOOPERATIVE
UNDESIRABLE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | DESIRABLE

MATURE
UNDESIRABLE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | DESIRABLE

IMMATURE
UNDESIRABLE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | DESIRABLE

16. The trait that I am to find out about is:

COOPERATIVE
NO | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | YES

UNCOOPERATIVE
NO | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | YES

MATURE
NO | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | YES

IMMATURE
NO | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | YES

17. Which of the following are true?

A. The instructions explicitly told me that I am to try and discover the respondent's standing on a trait.

FALSE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | TRUE

B. The instructions explicitly told me that I should be likable during the interview:

FALSE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | TRUE

C. The instructions explicitly told me that I will be meeting with the experimenter to discuss my interview questions:

FALSE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | TRUE

D. The instructions explicitly told me that my judgment will be compared with the respondent's score on a personality scale:

FALSE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | TRUE

E. The instructions explicitly told me that the respondent will not see my judgment:

FALSE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | TRUE

18. How important is it for you to do well when you make your trait judgment? That is, how important is it that your judgment matches the respondent's true standing on the trait?

IMPORTANT | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | UNIMPORTANT

19. How confident are you that your questions will permit you to make an accurate trait judgment?

CONFIDENT | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | UNCONFIDENT

20. How desirable are each of the following traits?

COOPERATIVE

UNDESIRABLE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | DESIRABLE

UNCOOPERATIVE

UNDESIRABLE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | DESIRABLE

MATURE

UNDESIRABLE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | DESIRABLE

IMMATURE

UNDESIRABLE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | DESIRABLE

Suppose 100 students were randomly selected from introductory psychology courses here at UF. How many of these students do you believe could be classified as:

A. Mature _____

B. Cooperative _____

C. Immature _____

D. Uncooperative _____

REFERENCES

Alicke, M. D. (1985). Global self-evaluations as determined by the desirability and controllability of trait adjectives. Journal of Personality and Social Psychology, 49, 1621-1630.

Alicke, M. D. (1987). Public explanation and private ratiocination: Communication between the public and private selves. In R. H. Hogan & W. H. Jones (Eds.), Perspectives in personality: Theory, measurement, and interpersonal dynamics (Vol. 2, pp. 143-180). Greenwich, CT: JAI Press.

Alison, S. T., & Messick, D. M. (1988). The feature-positive effect, attitude strength, and degree of perceived consensus. Personality and Social Psychology Bulletin, 14, 231-241.

Amabile, T. M. (1985). Motivation and creativity: Effects of motivational orientation on creative writers. Journal of Personality and Social Psychology, 48, 393-399.

Asch, S. E. (1946). Forming impressions of personality. Journal of Abnormal and Social Psychology, 41, 258-290.

Bandura, A. (1987). Social foundations of thought and action: A cognitive theory. Englewood Cliffs, NJ: Prentice Hall.

Bassok, M. & Trope, Y., (1984). People's strategies for testing hypotheses about another's personality: Confirmatory or diagnostic? Social Cognition, 2, 199-216.

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. Journal of Personality and Social Psychology, 46, 610-620.

Baumeister, R. F., Hutton, D. G., & Tice, D. M. (in press). Cognitive processes during deliberate self-presentation: How self-presenters alter and misinterpret the behavior of their interaction partners. Journal of Experimental Social Psychology.

Baumeister, R. F., & Steinhilder, A. (1984). Paradoxical effects of supportive audiences on performance under pressure: Homefield disadvantage in sports championships. Journal of Personality and Social Psychology, 36, 608-618.

Bersheid, E., & Walster, E. (1977). Interpersonal Attraction. (2nd ed). Reading, MA: Addison-Wesley.

Carver, C. S., & Scheier, M. R. (1985). Aspects of self and the control of behavior. In B. R. Schlenker (Ed.), The self and social life (pp. 146-174). New York: McGraw-Hill.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. Journal of Personality and Social Psychology, 39, 752-766.

Cronbach, L. J. (1955). Processes affecting scores on "understanding of others" and "assumed similarity." Psychological Bulletin, 52, 177-193.

Deci, E. L., & Ryan, R. M. (1980). The empirical exploration of intrinsic motivational processes. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 13, pp. 39-80). New York: Academic Press.

DePaulo, B. M., Kenny, D. A., Hoover, C. W., Webb, W., & Oliver, P. V. (1987). Accuracy of person perception: Do people know what kinds of impressions they convey? Journal of Personality and Social Psychology, 52, 303-315.

DePaulo, B. M., Stone, J. I., & Lassiter, G. D. (1985). Deceiving and detecting deceit. In B. R. Schlenker (Ed.), The self and social life (pp. 323-370). New York: McGraw-Hill.

Doherty, K., Weigold, M. F., & Schlenker, B. R. (1989a). Accountability and performance: The mediating role of outcome expectations. Paper presented at the meeting of the American Psychological Association, New Orleans.

Doherty, K., Weigold, M. F., & Schlenker, B. R. (1989b). Accountability and affective associates of success and failure. Paper presented and the meeting of the Southeast Psychological Association, Washington.

Doherty, K., Weigold, M. F., & Schlenker, B. R. (in press). Self-serving interpretations of motives. *Personality and Social Psychology Bulletin*.

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.

Fischoff, B., & Beyeth-Marom, R. (1983). Hypothesis evaluations from a Bayesian perspective. *Psychological Review*, *90*, 239-260.

Ford, J. K., & Weldon E. (1981). Forewarning and accountability: Effects on memory-based interpersonal judgments. *Personality and Social Psychology Bulletin*, *7*, 264-268.

Frey, D. (1978). Reactions to success and failure in public and private conditions. *Journal of Experimental Social Psychology*, *17*, 172-179.

Frey, D. (1981a). Reversible and irreversible decisions: Preference for consonant information as a function of attractiveness of decision alternatives. *Personality and Social Psychology Bulletin*, *7*, 621-626.

Frey, D. (1981b). The effect of negative information about oneself and cost of information on preferences for information about the source of this feedback. *Journal of Experimental Social Psychology*, *17*, 621-626.

Frey, D. (1986). Recent research on selective exposure to information. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 19, pp. 41-80). New York: Academic Press.

Frey, D., & Rosch, M. (1984). Information seeking after decisions: The roles of novelty of information and decision reversibility. *Personality and Social Psychology Bulletin*, *10*, 91-98.

Funder, D. C. (1986). *Errors and mistakes: Evaluating the accuracy of social judgment*. Unpublished manuscript.

Gilbert, D. T., Pehlam, B. W., & Krum, D. S. (1988). On cognitive busyness: When person perceivers meet persons perceived. *Journal of Personality and Social Psychology*, *54*, 733-740.

Goffman, E. (1967). *Interaction ritual: Essays on face-to-face behavior*. New York: Pantheon.

Golub, H. F., Rossman, B. B., & Abelson, R. P. (1973). Social inference as a function of the number of instances and consistency of information presented. Journal of Personality and Social Psychology, 27, 19-33.

Greenwald, A. G., & Breckler, S. J. (1985). To whom is the self presented? In B. R. Schlenker (Ed.), The self and social life. New York: McGraw-Hill.

Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. British Journal of Psychology, 73, 407-420.

Griggs, R. A., & Cox, J. R. (1983). The effects of problem content and negation on Wason's selection task. Quarterly Journal of Experimental Psychology, 35a, 519-533.

Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. Psychological Review, 93, 258-268.

Hastie, R., & Rasinski, K. A. (1988). The concept of accuracy in social judgment. In D. Bar-tal & A. W. Kruglanski (Eds.), The social psychology of knowledge (pp. 193-208). Cambridge: Cambridge University Press.

Higgins, E. T., & Stangor, C. (1988). Context-driven social judgment and memory: When "behavior engulfs the field" in reconstructive memory. In D. Bar-tal & A. W. Kruglanski (Eds.), The social psychology of knowledge (pp. 262-297). Cambridge: Cambridge University Press.

Holyoak, K. J., & Gordon, P. C. (1984). Information processing and social cognition. In R. S. Wyer & T. K. Srull (Eds.), Handbook of social cognition (pp. 39-70). Hillsdale, NJ: Erlbaum.

Homans, G. C. (1961). Social behavior: Its elementary forms. New York: Harcourt, Brace & World.

Hovland, C. I., & Weis, W. (1953). Transmission of information concerning concepts through positive and negative instances. Journal of Experimental Psychology, 45, 175-182.

Janis, I. L. & Mann, L. (1977). Decision making: A psychological analysis of conflict, choice, and commitment. New York: Free Press.

Jenkins, H. M., & Ward, W. C. (1965). Judgments of contingency between responses and outcomes. Psychological Monographs, 79(1, Whole No. 594).

Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. Journal of Experimental Social Psychology, 3, 1-24.

Kelly, G. (1955). A theory of personality: The psychology of personal constructs. New York: Norton.

Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis testing. Psychological Review, 94, 211-228.

Kruglanski, A. W. (1988). Knowledge as a social psychological construct. In D. Bar-tal & A. W. Kruglanski (Eds.), Social psychology of knowledge. Cambridge: Cambridge University Press.

Kruglanski, A. W. (in press). The psychology of being "right": On the problem of accuracy in social perception and cognition. Psychological Bulletin.

Kruglanski, A. W., & Klar, K. (1985). Hypothesis testing. Social Cognition, 3, pp. 40-53.

Kruglanski, A. W., & Mayseless, O. (1988). Contextual effects in hypothesis testing: Effects of competing hypotheses and epistemic motivations. Social Cognition, 6, pp. 1-20.

Lingle, J. H., Altom, M. W., & Medin, D. L. (1984). Of cabbages and kings: Assessing the extendibility of natural object concept models to social things. In R. S. Wyer & T. K. Srull (Eds.), Handbook of social cognition (Vol. 1, pp. 39-70). Hillsdale, NJ: Erlbaum.

McAllister, D. W., Mitchell, T. R., & Beach, L. R. (1979). The contingency model for the selection of decision strategies: An empirical test of the effects of significance, accountability, and reversibility. Organizational Behavior and Human Performance, 24, 228-244.

Mettee, D. R., & Aronson, E. (1974). Affective reactions to appraisal from others. In T. L. Huston (Ed.), Foundations of interpersonal attraction (pp. 235-283). New York: Academic Press.

Platt, J. R. (1964). Strong inference. Science, 146, 347-353.

Pyszczynski, T. A., & Greenberg, J. (1981). The role of disconfirmed expectancies in the instigation of attributional processing. Journal of Personality and Social Psychology, 40, 31-38.

Pyszczynksi, T., & Greenberg, J. (1987). Toward an integration of cognitive and motivational perspectives on social inference: A biased hypothesis testing model. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 20, pp. 297-340). New York: Academic Press.

Pyszczynksi, T., Greenberg, J., & Holt, K. (1985). Maintaining consistency between self-serving beliefs and available data: A bias in information evaluation following success and failure. Personality and Social Psychology Bulletin, 11, 179-190.

Pyszczynski, T., Greenberg, J., & LaPrelle, J. (1985). Social comparison after success and failure: Biased search for information consistent with a self-serving conclusion. Journal of Experimental Social Psychology, 21, 195-211.

Rosenthal, R. (1971). Teacher expectations and their effects upon children. In G. S. Lesser (Ed.), Psychology and educational practice. Glenview, IL: Scott, Foresman.

Ross, L. (1977). The intuitive scientist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.) Advances in experimental social psychology (Vol. 10, pp. 173-220). New York: Academic Press.

Ross, L., Amabile, T. M., & Steinmentz, J. L. (1977). Social roles, social control, and biases in social perception processes. Journal of Personality and Social Psychology, 35, 485-494.

Rozelle, R. M., & Baxter, J. C. (1981). Influence of role pressures on the perceiver: Judgments of videotaped interviews varying judge accountability and responsibility. Journal of Applied Psychology, 66, 437-441.

Schlenker, B. R. (1980). Impression management: The self concept, social identity, and interpersonal relations. Melbourne, FL: Krieger.

Schlenker, B. R. (1982). Translating actions into attitudes: An identity-analytic approach to the explanation of social conduct. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 15, pp. 193-247). New York: Academic Press.

Schlenker, B. R. (1984). Identities, identifications, and relationships. In V. Derlega (Ed.), Communication, intimacy, and close relationships (pp. 71-104). New York: Academic Press.

Schlenker, B. R. (1985). Identity and self-identification. In B. R. Schlenker (Ed.), The self and social life (pp. 65-99). New York: McGraw-Hill.

Schlenker, B. R. (1986a). Self-identification: Toward an integration of the private and public self. In R. F. Baumeister (Ed.), Public self and private self (pp. 21-62). New York: Springer-Verlag.

Schlenker, B. R. (1986b). Personal accountability: Challenges and impediments in the quest for excellence. Technical Report for the Navy Personnel Research and Developments Center, San Diego.

Schlenker, B. R. (1987). Threats to identity: Self-identification and social stress. In C. R. Snyder & C. Ford (Eds.), Clinical and social psychological perspectives on negative life events (pp. 273-321). New York: Academic Press.

Schlenker, B. R., Hallam, J. R., & McCown, N. E. (1983). Motives and social evaluation: Actor-observer differences in the delineation of motives for a beneficial act. Journal of Experimental Social Psychology, 19, 254-273.

Schlenker, B. R., & Leary, M. R. (1982). Audiences reactions to self-enhancing, self-denigrating, and accurate self-presentations. Journal of Experimental Social Psychology, 18, 89-104.

Schlenker, B. R., & Weigold, M. F. (1989). Goals and the self-identification process: Creating desired identities. In L. Pervin (Ed.), Goal concepts in personality and social psychology (pp. 243-290). Hillsdale, NJ: Erlbaum.

Schlenker, B. R., & Weigold, M. F. (in press). Self-identification and accountability. In R. A. Giacalone and P. Rosenfeld (Eds.), Impression management in the organization. Hillsdale, NJ: Erlbaum.

Schlenker, B. R., Weigold, M. F., & Doherty, K. (in press). Coping with accountability: Self-identifications and evaluative reckonings. In C. R. Snyder & D. R. Forsyth (Eds.), The handbook of clinical and social psychology. New York: Pergamon.

Shrauger, J. S. (1975). Responses to evaluations as a function of initial self-perceptions. Psychological Bulletin, 82, 581-596.

Skov, R. B., & Sherman, S. J. (1986). Information-gathering processes: Diagnosticity, hypothesis confirmatory strategies and perceived hypothesis confirmation. Journal of Experimental Social Psychology, 22, 93-121.

Snyder, M. (1981). Seek and ye shall find: Testing hypotheses about other people. In E. T. Higgins, C. P. Heiman, & M. P. Zanna (Eds.), Social cognition: The Ontario symposium on personality and social psychology (pp. 277-304). Hillsdale, NJ: Erlbaum.

Snyder, M. (1984). When belief creates reality. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 18, pp. 247-305). New York: Academic Press.

Snyder, M. & Campbell, (1980). Testing hypotheses about other people: The role of the hypothesis. Personality and Social Psychology Bulletin, 6, 421-426.

Snyder, M. & Cantor, N. (1979). Testing hypotheses about other people: The use of historical knowledge. Journal of Experimental Social Psychology, 15, 330-342.

Snyder, M. & Gangstead, S. (1981). Hypothesis testing processes. In J. H. Harvey, W. J. Ickes, & R. F. Kidd (Eds.), New directions in attribution research (Vol. 3, pp. 171-195). Hillsdale, NJ: Erlbaum.

Snyder, M. & Swann, W. B., (1978). Hypothesis testing in social interaction. Journal of Personality and Social Psychology, 36, 1202-1212.

Swann, W. B., Jr. (1983). Self-verification: Bringing social reality into harmony with the self. In J. Suls & A. G. Greenwald (Eds.), Psychological perspectives on the self (Vol. 2, pp. 33-66). Hillsdale, NJ: Erlbaum.

Swann, W. B., Jr. (1984). Quest for accuracy in person perception: A matter of pragmatics. Psychological Review, 91, 457-477.

Swann, W. B., Jr. (1985). The self as architect of social reality. In B. R. Schlenker (Ed.), The self and social life (pp. 100-125). New York: McGraw-Hill.

Swann, W. B., & Guliano, T. (1987). Confirmatory search strategies in social interaction: What, when, why, and with what consequences. Journal of Social and Clinical Psychology, 5, 511-524.

Swann, W. B., Griffin, J. J., Predmore, S. C., & Gaines, B. (1987). The cognitive-affective crossfire: When self-consistency confronts self-enhancement. Journal of Personality and Social Psychology, 52, 881-889.

Tetlock, P. E. (1983a). Accountability and the complexity of thought. Journal of Personality and Social Psychology, 45, 74-83.

Tetlock, P. E. (1983b). Accountability and the perseverance of first impressions. Social Psychology Quarterly, 46, 285-292.

Tetlock, P. E. (1985). Accountability: The neglected social context of judgment and thought. In B. M. Staw & L. Cummings (Eds.), Research in Organizational Behavior (Vol. 9, pp. 279-332). Greenwich, CT: JAI Press.

Tetlock, P. E., & Boettger, R. (in press). Accountability: A social magnifier of the dilution effect. Journal of Personality and Social Psychology.

Tetlock, P. E., & Kim, J. I. (1987). Accountability and judgment processes in a personality prediction task. Journal of Personality and Social Psychology, 52, 700-709.

Tetlock, P. E., Skitka, L., & Boettger, R. (in press). Social and cognitive strategies for coping with accountability: Conformity, complexity, and bolstering. Journal of Personality and Social Psychology.

Trope, Y. (1983). Self-assessment in achievement behavior. In J. Suls & A. G. Greenwald (Eds.), Psychological perspectives on the self (Vol. 2, pp. 93-121). Hillsdale, NJ: Erlbaum.

Trope, Y. (1986). Self-enhancement and self-assessment in achievement behavior. In R. M. Sorrentino & E. T. Higgins (Eds.), Handbook of motivation and cognition (pp. 350-378). New York: Guilford.

Trope, Y. & Bassok, M. (1982). Confirmatory and diagnosing strategies in social information gathering. Journal of Personality and Social Psychology, 43, 22-34.

Trope, Y. & Bassok, (1984). Information gathering strategies in hypothesis testing. Journal of Experimental Social Psychology, 19, 560-576.

Trope, Y., Bassok, M., & Alon, E. (1984). The questions lay interviewers ask. Journal of Personality, 52, 90-106.

Tversky, A. (1977). Features of similarity. Psychological Review, 84, 327-352.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. Quarterly Journal of Experimental Psychology, 12, 129-140.

Wason, P. C. (1968). On the failure to eliminate hypotheses: A second look. In P. C. Wason & P. N. Johnson-Laird (Eds.), Thinking and reasoning (pp. 165-174). Harmondsworth, Middlesex, England: Penguin.

Wason, P. C., & Johnson-Laird, P. N. (1972). Psychology of reasoning: Structure and content. London: Batsford.

Weigold, M. F., & Schlenker, B. R. (1989). Accountability and risk taking. Unpublished manuscript.

Wicklund, R. A. (1975). Objective self-awareness. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 8, pp. 233-275). New York: Academic Press.

BIOGRAPHICAL SKETCH

Michael Weigold was born in New York City in 1958, the first child of Frederick and Lorraine Weigold. His parents were both born in New York. Frederick Weigold was a firefighter in New York City until his retirement. Lorraine Weigold was a homemaker and is a clerk in the municipal government of Clarkstown, New York. Weigold completed his undergraduate studies in English at the State University of New York at Albany in 1980. After a stint in retail management, he returned to school and received a master's degree in journalism and mass communications in 1985. That same year he began graduate study in social psychology. His research, conducted under the guidance of Dr. Barry Schlenker, was in the areas of accountability and self-presentation. In 1989 Weigold accepted a position as Assistant Professor in Advertising in the College of Journalism and Communications at the University of Florida. He is married to Mary Ann Ferguson and is the father of a son, Sean.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Barry Schlenker, Chairman
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Mark Alicke
Assistant Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Lawrence J. Severy

Lawrence Severy
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

John D Lynch

John Lynch
Associate Professor of Marketing

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Richard Griggs

Richard Griggs
Professor of Psychology

This dissertation was submitted to the Graduate Faculty of the Department of Psychology in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1989

Dean, Graduate School